An Investigation of Finding and Refinding
Information on the Web

Robert G. Capra III

Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Computer Science and Applications

Dr. Manuel A. Pérez-Quiñones, Chairman
Dr. Dan Dunlap
Dr. H. Rex Hartson
Dr. David Novick
Dr. Naren Ramakrishnan

February 16, 2006
Blacksburg, Virginia

Keywords: information refinding, information science, personal information management

An Investigation of Finding and Refinding
Information on the Web

Robert G. Capra III

Abstract

Refinding electronic information is a common problem, yet it has received less study than
the problem of how to find information for the first time. In this dissertation, I examine
how people approach tasks to refind information they have seen on the Web and factors
that may affect refinding. I conducted a controlled, laboratory study in which
participants participated in two sessions: one to find information for a set of 18 tasks and
a second session, about a week later, to attempt a set of counterpart refinding tasks.
Results indicate that finding and refinding do have differences, but not for all types of
tasks. The use of Web search engines was not observed to change significantly from the
first session to the second. However, for tasks that participants were more familiar with,
search engines were used less. Tasks that involved refinding a subset of the information
that was found in the first session took longer to complete and were perceived as more
difficult. Participants often went directly to known resources on the Web to start their
searches. These sets of known resources included many on-line counterparts of paper
resources such as telephone directories, dictionaries, and newspapers. For many tasks,
participants used the same starting strategy to refind the information that they used find
it, indicating strong patterns of access. This work contributes to the base of knowledge
about how people refind information and the factors that affect refinding. It also
contributes to the research field of information refinding and personal information
management by identifying dimensions and factors that affect refinding. The results
reported have implications for the design of Web sites and information repositories, the
design of tools to help users find and refind information, and for the research community
studying personal information management.

To my family

Acknowledgements

First and foremost, I thank God for being a constant guide in my life and for helping me to accomplish the things that I am supposed to accomplish.

I have the most wonderful wife and family that anyone could ask for and they have supported me tirelessly throughout graduate school. To Miranda, I thank you for all you do for me. I cannot imagine a more loving, understanding, fun, and supportive wife. To my parents Bob and Carole, thanks for always giving me the love and confidence I needed. To my brother Tony, thanks for all the encouragement over late-night instant messages and for all your support. I have also been fortunate to have wonderful grandparents and a warm and supporting set of in-laws who all deserve special thanks.

Very special thanks go to my advisor, Dr. Manuel Pérez-Quiñones. He has helped shape my research and taught me much about being a mentor, advisor, and friend. I also thank my committee members: Dr. Naren Ramakrishnan for many enjoyable and fruitful "hallway" discussions, Dr. Rex Hartson for his friendship and deep knowledge of HCI, Dr. David Novick for his insights and advice on structuring the dissertation, and Dr. Dan Dunlap for his viewpoints and philosophical questions. Dr. Bob Williges also provided me with much valuable advice, especially statistics assistance, and guidance over the years and deserves special thanks. Thanks to Dr. Cal Ribbens for the many opportunities and good advice he has given me. Thanks to Mary Pinney for assistance running the main study, and to Miranda and Con for additional help with the study and survey. I also thank all the people that participated in my study and responded to my survey. Thanks to Dr. Andrea Kavanaugh and Dr. Francis Quek for allowing me to have a great working environment at the Center for HCI. At different points in my time at Virginia Tech, I have been supported by funding from the NSF, IBM, and the Virginia Tech Department of Computer Science and I thank them all.

Many people at Virginia Tech have been great friends and colleagues during the course of my studies. Nathan, Tanya, Wendy, Chris, Craig, and Jason are all great friends who also enjoy a good discussion about current HCI and computer topics and made the graduate process more enjoyable. Thanks to Sammy for blazing the trail and for many valuable discussions about research and life. Thanks also go to the Dialogue Research Group and Center for HCI folks: Con, Kibum, Manas, Jon, Pardha, and Jai in particular have been great colleagues, been willing to discuss my research, and generally made all those hours in 3160 more enjoyable.

Yahoo Music, Jimmy John's, Au Bon Pain, Netflix, and many carbonated beverages also helped make this dissertation possible ;-)

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The problem of *refinding* information on the Web has received less study than the problem of *finding* information, despite indications that refinding and organizing Web information for re-use are problems for many users [GVU98] [JBD01]. Research about Web information finding has been conducted in the fields of human-computer interaction, information retrieval, and information science. However, only a few research projects have specifically investigated refinding [AJK05] [BS04] [TAA+04] [JBD03] [MB97a]. Of these, only one study directly compared finding and refinding in an experimental setting [MB97a]. The focus of my Ph.D. dissertation is to examine differences between finding and refinding and to identify factors that affect the refinding process.

1.1 Motivation and Problem Statement

Looking for information on the Web has become a preferred method for satisfying many information needs. On a typical day, an estimated 94 million Americans use the Internet and 60 million use search engines [Pew05a][RS05]. Users also manage and search other types of electronic information in the form of email messages, documents, images, and attachments. In addition, users often have multiple computing devices: a computer at home, a computer at work, a cell phone, a PDA. With all this electronic information, studies have documented the feeling many users have of "information overload." Of increasing interest is how users organize, manage, and refind electronic information – an area of research that is becoming known as personal information management (PIM). Studies have documented that organizing and refinding information can be a problem for many users [JDO01][GVU98][WH01]. PIM recognizes that people have a variety of needs to store, re-access, and reuse information that need to be supported by information sources, tools, and interfaces. Furthermore, the PIM research community has recognized that little research has been conducted to understand users' methods and strategies for organizing and refinding information. It is the combination of the lack of research in understanding refinding and the need by users for better tools for refinding that are large motivators behind my research. In this dissertation, I specifically focus on investigating aspects of refinding information on the Web.

1.2 Research Goals

The high-level goals of my dissertation research are:

- To gain a better understanding of the process of refinding information on the Web and the factors that affect it.

- To examine the differences between finding and refinding information on the Web. To investigate if/how refinding is a different process from finding.

- To explore which tasks and types of information are easy to refind on the Web and which are difficult and what factors affect difficulty.

Underlying each of these goals is an overall motivation to gain insights that can be used to improve tools and interface designs for Web information refinding.

1.3   Research Questions

I explore four main research questions in this dissertation as refinements of the high-level goals. Each of these is described briefly below. Specific hypotheses are described in Chapter 4, Method.

1.3.1   RQ1 – Differences between finding and refinding

RQ1 – Do finding and refinding differ? Specifically, are there measurable differences between finding and refinding in terms of user knowledge/familiarity, perceived difficulty, search engine use, and task length?

Redoing a task a second time is often easier than performing it for the first time and may employ different processes and skills (consider refinding an office or building in an unfamiliar area of town). Similarly, I anticipate that refinding information on the Web differs from the process of finding information for the first time. Finding information on the Web is commonly an exploratory process with uncertainty about if or where information may be found. In contrast, refinding is typically a more directed, constrained process. While some finding tasks may be directed, refinding can rely on cues and associations from the finding experience, while finding can only rely on knowledge of similar tasks and from more general experiences.

Research from cognitive psychology has shown that associations, cues, and information encoding are important aspects in retrieving items from memory [EH89] [And80]. Experience and frequency of exposure affect recall and speed of recall [And80, p. 172]. Because refinding can rely on knowledge of the prior finding experience, it is expected that finding and refinding will differ in terms of the amount of knowledge/certainty the user has about the location of the information, the perceived difficulty of the task, and the amount of time it will take to find or refind the information.

Prior research has not specifically looked specifically at the differences between finding and refinding of Web information, except [MB97a]. Previous work has investigated how people organize information and keep it for reuse [JBD03] and has developed and tested tools to assist users in the organization and refinding of electronic information including Web information. However, it has not focused on the processes and characteristics of refinding that are examined here.

The contributions of my work show that finding and refinding have differences, but differences were not observed for all tasks. Differences were found for some types of tasks but not for others. Task familiarity ratings were generally higher for refinding tasks and the perceived difficulty of a task can have effects on finding and refinding behaviors. However, no significant change in Web search engine use was observed between an initial finding session and a subsequent refinding session. These results have implications for the design of Web sites, refinding tools, Web search engines, and additional studies of refinding. Tools and search engines could use these findings to help identify when a person is trying to refind information and to provide additional cues and integrate previously viewed results.

1.3.2   RQ2 – Task Differences

RQ2 – Does task and familiarity with a task affect finding and refinding? What types of tasks are easy and which are difficult? Specifically, how do task and task familiarity (knowledge of how to perform the task) affect user knowledge/familiarity, perceived difficulty, search engine use, and task length?

Task type/task domain is commonly an important factor in human-computer interaction and has been found to be a significant factor in studies of Web information finding [WRM+02] [TFK+03]. Taxonomic

classification of Web behaviors is an important aspect of studying and describing Web usage and several taxonomies of tasks performed on the Web have been proposed [BJW+99] [MPC01] [Bro02]. It is anticipated that task type will have a significant effect on Web information finding and refinding.

The level of familiarity that a searcher has with the domain being searched is a topic of research within the information science community [Wil04] [KC02] [SR03] [BB02] [MS98]. Domain knowledge is generally thought to have an effect on information seeking behaviors, but the question has not been completely resolved as to the specific effects [Wil04]. Refinding tasks differ from those typically studied by information science in that refinding deals with information that has been seen before. However, factors that affect finding may also affect refinding. As such, I have chosen to look at the possible effects of a related factor – task familiarity. Whereas domain knowledge deals with the searcher's familiarity with the general topic or domain being searched (sometimes called *topic familiarity*), the concept of *task familiarity* is concerned with the searcher's knowledge of a specific task or set of related tasks. It is expected that tasks for which a person has high task familiarity will be completed more easily than other tasks.

My work contributes to this research question by showing that task is an important factor for both finding and refinding tasks. Furthermore, differences between sessions were not observed for all tasks (i.e. there was an interaction effect of Session x Task). These results are consistent with results from information science literature in which task was found to be an important factor to consider when studying information seeking behaviors [WRM+02] [TFK+03]. In my research, tasks were found to fall into three general categories: fairly *easy* tasks which participants were familiar with, *difficult* tasks that involved several steps or topics with which the participants were unfamiliar, and a "*medium*" set of tasks that were in-between the easy and difficult sets. Task affected the use of Web search engines. Prior familiarity with a task also had significant effects (e.g. high prior familiarity tasks had lower search engine use).

1.3.3   RQ3 – Finding to Refinding Relationship

RQ3 – How does the relationship of the refinding task to the finding task affect the refinding process? Specifically, does it affect user knowledge/familiarity, perceived difficulty, search engine use, and task length?

The relationship – or similarity – of a finding task to a counterpart refinding task may have a strong affect on the refinding process. At first glance, we might expect that if *the exact same task* were given first as a finding task and later as a refinding task, that it would be easier the second time. An effect known as the encoding-specificity principle from cognitive psychology lends some support to this idea: "The probability of recalling an item at test depends on the similarity of its encoding at test and its original encoding at study" (from Tulving, as cited in [And80, p. 212]).

However, refinding does not always manifest itself as exactly the same task as finding. People often attempt to "refind" other, related information that was seen along a *path* traced while finding something else, or they may look for a *subset* of the information originally found. Additionally, information on the Web is dynamic and may be *moved* from the location it was originally found at, or may be deleted altogether [Spi03]. This research question addresses these four types of refinding tasks as they relate to the original finding task: exactly the same (E), on the same path (P), subset (S), and moved (M).

My work makes several significant contributions to this area. The study of refinding as a distinct process is so new that the distinctions of exact, path, and subset are contributions to the study of refinding. Results of my research indicate that refinding moved or deleted items can be so difficult that users may abandon the attempt, or not try at all. Refinding an item from a subset task took longer and was perceived as more difficult than exact or path refinding tasks. Subset tasks are an aspect of refinding that should be studied further. My results on tasks to refind information on the same path were possibly affected by an unintentional bias in the task grouping that was not realized until the analysis stage. However, it is unclear if there was a bias, or if path tasks may orient toward resources that are more familiar and require less use

of Web search engines. This may be because they often use resources that contain sets of valuable information and are well known.

1.3.4   RQ4 – Task Details and Strategies

RQ4 – What strategies are used in finding and refinding? Are there observable patterns of strategy use across types of tasks or from finding to refinding for a single task?

Information science has observed strategies that people use when conducting research and finding information. These observations have been codified into models of information seeking and information searching behaviors [Wil99]. For instance, as described by Wilson [Wil99], Ellis outlined features of information seeking (starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending) and Kuhlthau [Khu91] described stages in the information seeking process (initiation, selection, exploration, formulation, collection, and presentation). Many of these models focus on the broad process of performing research and information seeking for evolving, dynamic information tasks.

Personal information management and information refinding are often focused on well-defined, directed tasks to find or refind specific pieces of information such as a phone number, Web page, or specific document. These tasks differ from the broad information seeking tasks considered in many existing models of finding. A better understanding of the factors and strategies for personal information management and refinding are needed to develop models that can help inform the design of information and tools to support users' refinding needs.

To address this research question, I consider strategies that participants used to find and refind information as part of the main study of this dissertation. Previous research questions in this dissertation focused mainly on quantitative measures and analysis of controlled factors. For this research question, in order to better understand the strategies used, a detailed task-by-task analysis was conducted. This analysis examined the specific starting strategy used by each participant in each task and the resource used to locate the information.

My results indicate that users do have patterns that are followed for many types of tasks and that users are well versed in a (small) set of Internet resources (i.e. they know the URL for these sites and can type them directly into a Web browser to teleport to the site). Many of these known resources are Internet counterparts of longstanding real-world resources such as dictionaries, phone books, newspapers, and maps. In addition, users often used the same starting strategies to refind a piece of information that they did to find it. In cases where they deviated, they often used shortcuts or waypoints learned from the finding process to simplify the refinding. In addition, often users ended their refinding at the same resource they used when finding. However, in some cases, the location of the information or the resource used did not matter and the information was "refound" from a different resource than it was found. This is an example of the fungibility of resources to provide information.

1.4   Reader's Guide

This dissertation is organized in the following way. This chapter presents an introduction to the topic and focus of my dissertation with a summary of the major research questions, motivations, and results. Chapter 2 provides background on finding and refinding research, along with an extensive review of tools to support finding, refinding, and information reuse. The literature discussed is from several related areas including information science, computer science, cognitive science, and library science.

Chapter 3 describes a survey that I conducted to investigate what types of tasks people do frequently and infrequently on the Web. The results of this survey are presented here and were used to help generate the tasks for the main refinding study of finding and refinding presented in the dissertation. Chapter 4 outlines the method used for the main study of this dissertation. This experiment studied finding and refinding

behaviors of 17 participants on a set of 18 tasks to locate information on the Web. Chapter 4 also presents the specific hypotheses to be examined and relates the hypotheses to the main research questions of this dissertation.

Chapter 5 describes the results of analysis of a large number of dependent measures collected. The analysis techniques are described and results of the statistical analyses are presented. The results are organized around the specific analyses conducted. Chapter 6 discusses and summarizes the results of the analyses, presented as they relate to each research question. For each research question, the main findings are presented along with interesting observations and implications of the results.

Chapter 7 outlines a plan for future exploration of personal information management tasks, including further study of finding and refinding. Promising directions for future research are described and several specific studies are outlined.

This page intentionally left blank.

Chapter 2

Introduction to Finding and Refinding Research

Increasingly, organizing and managing electronic information is a problem for many computer users. Several recent research efforts are focusing on understanding not just how people search for and save information when seeking it the first time, but also on how they subsequently try to refind and reuse the information.

Many factors play a role in how people search for information. Studies from research areas including information retrieval, digital libraries, information science, cognitive science, and human-computer interaction have explored differences based on task type and task domain, individual differences, searcher experience, searcher familiarity with the topic, and interface/system differences.

Refinding is a fairly new area of study, with little prior research that has focused directly on the factors that affect refinding, the strategies that people use, and the differences between finding and refinding. However, a number of research projects, studies, and tools have addressed issues related to refinding and personal information management. In this section, a brief introduction to refinding is presented, followed by sections that describe important aspects of finding, refinding, and tools to support PIM.

Several major factors are important to the study of both PIM and refinding and will be reviewed in the following sections. The task being performed and the topic domain can have a large effect on how information is found, organized, and reused – these aspects are discussed in section 2.2. Individual differences of the person performing the search can also have large effects. Individual factors such as domain knowledge and expertise, knowledge of the task, and cognitive aspects are reviewed in section 2.3. Although it has not been studied extensively, the amount of time that passes before information is needed to be refound may have effects on the approaches used to refind it. In addition, assessing the future value of information is an especially difficult task for users. These aspects are considered in section 2.4. Information science has examined factors that affect the finding process and codified models of information seeking behaviors. While refinding does have differences with finding, these models are important to help understand the relationship of the finding and refinding processes and also provide essential starting points for developing models of refinding. Section 2.5 describes several of these information seeking models in more detail. The remaining sections in this chapter describe previous studies of information refinding and reuse and tools to help support users' PIM.

2.1 Introduction

The essence of refinding is: something was found, is needed again, and must be located.

Refinding can take many forms, from trying to relocate a key chain to trying to refind an email message that contained important information. In this dissertation, the focus is on refinding electronic information – more specifically, it is on refinding information that has been seen previously on the World Wide Web (WWW, or Web, for short). Refinding electronic information is a large component of personal information management (PIM). PIM is a broad area that considers how people organize, manage, archive, refind, and reuse a variety of types of information including paper documents, email messages, Web pages, electronic documents, files, contact and calendar information, and just about any other form of information.

Refinding and personal information management have links to several other research areas in addition to research on information seeking. These include the semantic Web, research on the use of contextual information in interfaces and information retrieval, mobile computing, ubiquitous computing, mobile access to information, and voice user interfaces to information.

2.1.1   Refinding is difficult to study

The nature of refinding requires that a person: 1) has found some information, and 2) has a need or motivation to relocate this previously found information. Finding, and especially refinding are often motivated by unpredictable, external events. The process of searching for information is often opportunistic and can evolve over time. These aspects make it difficult to observe refinding episodes as they occur and can be difficult to recreate in a laboratory setting. Interview and diary studies can investigate refinding by asking users to recall refinding events and critical incidents and have been adapted to fit the unpredictable nature of refinding needs (see [TAA+04] for an example of a modified diary technique). Studies of data logs (such as the Web logs studied in [TG97]) can provide interesting results of aggregate behaviors, but logs often do not provide the rich insights and details from direct observation or interviews. Studying refinding in a laboratory setting has the benefit of being able to better control external factors and to isolate factors of interest. However, laboratory studies can lose important aspects of real-world activity.

In the research presented here, I chose to study refinding using a laboratory study based on a set of tasks that were developed with the help of an empirical study of frequent and infrequent tasks. I chose a laboratory study in order to help focus on particular factors that might affect refinding.

2.1.2   Finding and refinding are related, but different tasks

Finding involves the searcher in a process of looking for information. The purpose and scope of a finding activity may be well defined (e.g., find Rob Capra's Web page), or may be more vague (e.g., find information about the history of Virginia Tech). Finding, by nature, is an exploratory process. The searcher is looking for something not found before and thus has some level of uncertainty about what (if anything) will be found and where it will be located. Certainly, there are some finding tasks for which users are confident of their outcome based on previous experience with similar tasks. However, many finding tasks do involve uncertainty, both in terms of where and if the desired information will be found, and how to go about searching for it.

Refinding also involves the searcher in the process of looking for information, but refinding is typically more purposeful and well defined. Often people try to refind specific items. By definition, refinding involves looking for something that has been seen before or is known to exist[1]. Refinding can make use of knowledge that is remembered from the finding process such as starting points and waypoints, and may rely on established patterns of access. In one study, researchers found that memories of patterns of access were so strong that some participants "remembered" steps in a Web information-seeking task that they had not actually performed (such as starting at a particular site or using a particular search engine) [MB97a]. However, very little research has directly looked at Web information refinding, and even less has examined specific differences between finding and refinding.

---

[1] An interesting question and point of debate about this definition of refinding is, "Can you refind something that you know to exist but have not ever seen, such as an email that someone sent you but you have not yet read?"

2.2   Task

Task/Task type is commonly an important factor in HCI and search behaviors. In this section, we review several dimensions of tasks that are important to finding and refinding.

2.2.1   Taxonomies

Taxonomic classification of Web behaviors is an important aspect of studying and describing Web usage. Several taxonomies of tasks performed on the Web have been proposed [BJW+99] [MPC01] [Bro02].

Byrne et al. [BJW+99] proposed a taxonomy of tasks at an operational level, close to that of a Web browser interface. The top-level of their taxonomy includes the categories: use information, locate on page, go to page, provide information, configure browser, and react to environment. These categories are refined with actions such as read, view, download, print, display, and duplicate. Byrne et al. conducted a naturalistic study of Web behaviors and applied their taxonomy to classify observed tasks.

The taxonomy that Bryne et al. [BJW+99] proposed is useful for examining the methods used to accomplish higher-level tasks and may be considered at the levels of methods and operators in GOMS [CMN83][JK96]. For example, in a task to find and print the current weather for Blacksburg, VA, a user might go to a page (www.noaa.gov), locate on a page (find a text box to enter a city or zip code), provide information (zip code = 24061), and configure the browser (for printing the page).

Broder [Bro02] identified three types of motivations for a Web search: navigational, informational, and transactional. Navigational searches have a goal of finding "a particular site that the user has in mind" [Bro02, p. 5] and "is sometimes referred as 'known item' search in classical IR" [Bro02, p.5]. Informational searches are traditional searches for information on a particular topic. These queries may be broad in nature or quite narrow. Transactional queries are performed to find a site where additional interaction will occur, such as shopping. Broder conducted two studies of searches on AltaVista and found that 20-25% of queries were Navigational, 40-50% were Informational, and 22-36% were Transactional. Using search for navigational purposes as noted by Broder [Bro02] is similar to the use of a search engine in specifying the addressability of a Web page as noted by Naren Ramakrishnan [Ram05].

Morrison, Pirolli, and Card proposed three separate taxonomies based on the purpose, method, and content of a Web search [MPC01, p. 164]:

- Purpose taxonomy: "What was the primary reason for the respondent's search?"
  - find (25%), understand (24%), compare/choose (51%)
- Method taxonomy: "How did the respondent find the information?"
  - collect (71%), find (25%), monitor (2%), explore (2%)
- Content taxonomy: "For what type of information is the responded [sic] searching"
  - product info and purchase (30%), medical (18%), people (13%), misc (8%), travel (7%), business (7%), education (6%), job search (6%), finance (4%), news (1%)

2.2.2   Domain

Task type and task domain are important factors in human-computer interaction and have been found to be a significant factor in studies of Web information finding [WRM+02] [TFK+03]. Note that task domain as considered in this section is related to, but different from the concept of an individual's knowledge of the domain (often called domain knowledge or topic familiarity). Domain knowledge is discussed in section 2.3.1.

The topic of a search task has been shown to play a role in search behaviors. Woodruff et al. [WRM+02] conducted a study to investigate three different methods of presenting summaries of search results: plain text, thumbnail images, and thumbnail images that were enhanced by highlighting and emphasizing search

terms. Their study included search tasks to find specific information in four categories: picture, homepage, e-commerce, and medical information. They found that task category had a significant interaction with summary type for measures of task completion time and number of pages viewed. In addition, task category had an effect on the number of "false alarm rate" – a measure of how often participants visited pages that did not lead to a correct answer to the task.

Toms et al. [TFK+03] examined search strategies in four domains (consumer health, shopping, travel, and general research) using a hybrid interface that included both a list of Google's top-level directory categories and a text box to type a search query to Google. They report differences between the categories – for the categories of shopping and travel, participants spent "more time exploring within a site" [TFK+03, p. 309].

2.2.3   Task Type and Task Complexity

Task complexity is itself a complex topic with several dimensions.

In information science, task complexity can be considered in terms of: repetitivity, *a priori* determinability, the number of alternative paths to the information, number of goals, uncertainty, number of inputs, and other factors [BJ95, p.193] [Bys02, p. 582]. Information science studies and models often consider the predeterminability of task outcomes as a defining characteristic of the complexity of a task [BJ95] [Vak99] [WBL+06].

Many of the tasks addressed in PIM are directed information seeking tasks – in other words, they are concerned with finding specific information, the structure of which is probably known to the searcher. Often these tasks are fairly simple in concept, although the information itself may be difficult to locate. For example, a user may know that she needs John Doe's phone number, but cannot find it. However, information is often found in the course of addressing more complex or open-ended tasks.

In HCI and WWW research, information-seeking behaviors are often considered as directed or undirected – searching or browsing. Searching is concerned with looking for a known item, while browsing is a process of looking to see what is available [JF97]. Searching and browsing have complimentary benefits – searching is good for finding pages with specific terms quickly and browsing is useful in situations where a good search query cannot be constructed (such as relativistic comparisons) [OC03]. For complex tasks both searching and browsing may be beneficial, although it can be difficult to combine the two approaches [OC03].

Task structure is a related concept that can also play an important role in how information is sought. Vakarri [Vak99, p.832] summarizes work by Belkin, Marchetti, and Cool; Ellis and Haugan; and Bates as follows: if an information-seeking problem is ill structured, the overall search strategy used will be that of browsing (scanning, browsing, learning, recognition, surveying-chaining, journal run). If the problem is structured, the strategy will be that of querying (selecting, specification, and ending). [Vak99, p.832].

Lieberman, Nardi, and Wright [LNW99] describe valuable "nuggets" of semi-structured information such as email addresses and phone numbers that are embedded in documents such as Web pages. These information nuggets are often the targets of PIM searches. The concept presented by Lieberman et al. is similar to the idea behind the semantic Web – that important and actionable pieces of information can be recognized and tagged with metadata for easy use, re-finding, and reuse.

Information refinding can be a challenging problem when the information is static, but information on the Web can be dynamic in terms of its location and lifetime. Studies have documented the dynamic nature of Web pages [Spi03] [BC00] [Koe02]. The ever-changing nature of the Web makes refinding even more difficult and is an important aspect for information refinding tools, interfaces, and research.

2.3   Individual

It is well known that individual differences affect many aspects of human-computer interaction and how individuals deal with information. A special issue of the Journal of the American Society for Information Science and Technology (April 2000) was devoted to "Individual differences in Virtual Environments", including not only virtual reality, but also electronic information spaces. Models of information seeking include components for individual factors [BJ95]. Factors such as cognitive style [FMM01], self-efficacy [FMM01], experience [LBW00] [JCW03] [HS00], and gender [FMM01] have been studied for their effect on information search and retrieval. Preliminary results of a recent study of how people use the Web indicate that there are large individual differences in knowledge and experience that affect strategies used and task completion time [Har02].

2.3.1   Domain Knowledge / Expertise

Topic familiarity refers to the searcher's general knowledge of the topic of the information search and is also referred to as domain knowledge. Familiarity with a topic or task is commonly regarded to be an important factor in information search and retrieval [KC02]. Wildemuth [Wil04] presents good review of prior research on the effects of domain knowledge on search tactics, searching behaviors, search term selection, retrieval success, and outcomes of searching. Based on her review, she concludes that, "the evidence concerning the effects of domain knowledge on searching behaviors is inconclusive," [Wil04, p.247] but also notes, "there is, however, consistent support for the propositions that differences in domain knowledge affect searchers' selection of terms and that databases can effectively augment personal knowledge in problem-solving situations" [Wil04, p. 248].

Kelly and Cool [KC02] examined the effect of topic familiarity on information search behaviors with an article database. They found evidence that familiarity affected the amount of time participants spend reading documents and the efficacy of the searches. They also found differences based on familiarity in the amount of time participants spent reading documents. Although these differences were not statistically significant in their study, they felt that the trend showed that "the more familiar one is with a search topic, the less time one spends reading" [KC02, p.75].

Bhavnani [Bha02] examined the effects of domain knowledge on searching in the areas of healthcare and shopping. Participants in Bhavnani's study were expert in one of the domains but not the other. The participants were asked to perform search tasks in both domains. The results showed that domain experts did exhibit different, more effective searching behaviors. Bhavnani noted that experts made use of domain knowledge to help their searches. He also noted that it might be difficult for novices to acquire the knowledge of experts by using search engines. In related work, Bhavnani et al. [BBJ+03], argue that, "domain portals… do provide reliable sources of information, but also do not provide the procedural knowledge to organize a search" [BBJ+03, p.394]. To address these issues, they developed a portal, Strategy Hub, designed to help users acquire this knowledge.

In an interesting study, Whittaker, Hirschberg, and Nakatani [WHN98] examined the effects of familiarity on searching an archive of recorded speech data. They found that "overall familiar material elicits more efficient search" [WHN98, p. 248]. They also looked at the effects of task type, finding that more complex task types resulted in less efficient searches than tasks directed at finding a single piece of information.

Zhang and Ghorbani [ZG04] describe the relationship between familiarity and trust that has been a topic of study in a variety of fields. They note that familiarity is a multi-dimensional concept and review five factors that may affect familiarity: "prior experience, repeated exposure, level of processing, study duration, and forgetting rate" [ZG04, p. 24]. They also note that measuring familiarity can be difficult because an individual's familiarity with a Web site may change over time and due to the factors that affect familiarity.

Often, expertise is considered along two dimensions: domain expertise and search expertise [HS00] [JCW03]. Domain experts are knowledgeable about the particular domain in which information is being sought. Search experts have specialized knowledge and experience with the tools or databases being used

in the search. For example, librarians may have knowledge about how to search particular databases, and Web experts may know special techniques for searching the Web.

Hölscher and Strube [HS00] report on an experimental comparison of these two types of expertise. A 2x2 study was conducted with Web expertise and domain knowledge as the independent factors. They found both "independent and combined effects" [HS00, p.345] of the two types of expertise. They report that participants with expertise in both areas did the best on search tasks. Participants with experience in one area used compensating strategies such as relying on terminology (domain experts). Participants with low expertise were "less flexible in their strategies" [HS00, p.345].

Jenkins, Corritore, and Wiedenbeck [JCW03] also conducted a study to examine domain expertise and Web expertise in a 2x2 study. Their results suggest that Web expertise is linked to search strategy such that Web novices search breadth-first and Web experts are more likely to investigate deeply into the results. Their results also suggest that domain experts evaluate search results more closely [JCW03].

Lazonder, Biemans, and Wopereis [LBW00] found that Web experts were better than Web novices at locating a Web site (searching), but that experts and novices performed about the same in locating information once the Web site had been found (browsing).

In a longitudinal study of Web use, Cothey [Cot02] observed that as users gained experience using the Web, they used the Web less often, used querying less and relied more on link-clicking, and used more diverse sets of Web hosts. Cothey notes that these were unexpected results. She discusses several possible confounding factors such as shifts in tasks and evolution of the Web, but gives reasons why these factors were unlikely to have been the cause of the observed results.

2.3.2   Task Knowledge

Whereas topic familiarity addresses how familiar a searcher is with a particular topic area, task knowledge measures knowledge about how to do a specific task. Two factors that contribute to task knowledge are examined in this dissertation: the *frequency* with which a person does a task, and the *familiarity* that a person has with where the information requested by the task is located. To the best of my knowledge, my work is the first to examine the effects of frequency and familiarity on refinding.

In a strict sense, information refinding by an individual can be viewed dichotomously – the information being sought has either been seen before and the current task is refinding, or this is the first time the information is being looked for and the task is finding. In a more relativistic sense, refinding can be considered in terms of frequency – how often does the searcher look for this information? For very frequent tasks, a user may have highly developed methods of access and may be more likely to "teleport" directly to the information. For medium frequency tasks, the refinding process may involve a mixture of recall and recognition. For very infrequent tasks, refinding may truly be "finding again".

2.4   Time

Elapsed time is known to be an important factor in forgetting [And80]. In addition to forgetting, additional stimuli may interfere with existing memories [And80], especially of similar items. Based on these effects studied by cognitive psychology, I expect that time (and the interference of similar items) plays an important role in information refinding. Specifically, the time that elapses from the last time that a piece of information was found to the time that it is needed to be refound is likely to be an important factor.

To the best of my knowledge, the specific effects of time on Web information refinding have not been examined extensively. However, particular time gaps have been used in studying other effects. For example, Jones and Dumais [JD86] conducted a study of the effects of spatial references to documents and used a 13-17 day gap between practice and recall sessions. Maglio and Barrett [MB97a] studied people's ability to recall and reproduce Web searches and used a gap of one day. In the main study presented in this dissertation, I used a time gap of approximately seven days.

Another factor related to time is how time affects the value of information. Information that is valuable right now may be less valuable in the future. Similarly, information that does not stand out as being important now may be realized to be very important at a later time. Wen [Wen03] conducted a study of what he refers to as "post-valued recall" (PVR) Web pages – pages "whose value is not recognized until some time after its initial retrieval." [Wen03, p. 185] In a study of PVR, Wen observed that users had difficulties relocating PVR pages and that they had more difficulty locating recalled pictures than recalled text, likely due to difficulties in using Web search engines to locate images.

Mary Pinney conducted study of PVR and refinding [CPP04] while working with the research group that I am a part of at Virginia Tech. In her study, she had two experimental sessions. In the first session, she informed half the participants that information found in the first session would need to be refound in the second session. The other half of the participants were not told specifically that they would be need to refind the information. In her results, she did not observe significant differences in refinding between the two groups. This result was not expected and may have been due to the difficulties in motivating PVR in a laboratory setting. Time and PVR are important aspects of refinding and personal information management and their effects should be explored in more detail.

2.5   Models

In this section, several models of information seeking are reviewed and their relation to information refinding and personal information management is discussed.

Models of information seeking include categorizations of task categories [BJ95], information types [BJ95], and stages of information seeking [Kuh91]. Byström and Järvelin propose a model that distinguishes tasks in terms of: the *a priori* determinability of the information needed, the process to obtain the information, and the task results. Tasks for which all three elements can be fully described in advance fall at one end of the spectrum (automatic information processing tasks) and are fairly well defined, while tasks that have many unknowns are at the other end (genuine decision tasks) and are more unstructured [BJ95]. Differences in how people approach the types of tasks have been shown [BJ95] [Bys02]. Byström and Järvelin also outline three types of information: problem information, domain information, and problem-solving information [BJ95]. Kuhlthau describes stages in the information seeking process: initiation, selection, exploration, formulation, collection, and presentation – and describes differences in thoughts, actions, and tasks for each stages [Kuh91].

2.5.1   Finding – Vakkari

In an effort to integrate research from information retrieval and information seeking, Vakkari [Vak99] outlined a model that connects task complexity, problem structure, and prior knowledge to information actions. In his model, shown in Figure 1, prior knowledge, task complexity, and problem structure all may influence each other, and each have effects on the information actions. He cites "both theoretical and empirical" [Vak99, p. 822] evidence to support the model, yet notes that additional research is needed to examine "the various relations between task complexity and the major dimensions of information activities." [Vak99, p835].

Prior knowledge
knowledge structures
• scope
• differentiation
• integration

Task complexity
predeterminability of
• info requirements
• process
• outcome

Information actions
• need and use
  of info and source types
• search strategies
• relevance assessment

Problem structure
• structured
• ill-structured
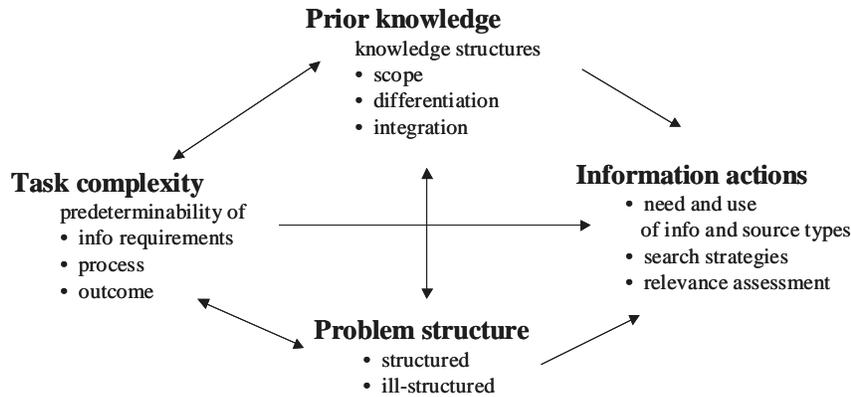
Figure 1. Model of Information Seeking from [Vak99, p.830][2]

Vakkari also notes that few studies have examined how prior knowledge relates to "the changes of cognitive structures of actors in information searching," [Vak99, p.828] and that "One of the main difficulties in research has been how to describe changes in understanding, i.e. in cognitive structures in a way that would connect the change process to changes in information actions" [Vak99, p.828].

In this dissertation, I do not attempt to examine cognitive structures of searchers and how they change, but I do focus on the role that prior knowledge and experience play in information seeking and present empirical results showing effects of prior knowledge on information actions. I present results from a controlled, laboratory study of finding and refinding information on the Web that examined the effects of task frequency and task familiarity (prior knowledge) on information seeking behaviors. My study focused on tasks with well-defined information goals.

2.5.2   Finding – Belkin

Belkin, Marchetti, and Cool [BMC93] describe a model of information seeking that characterizes interactions along a set of four dimensions:

Method of Interaction:	scanning to searching
Goal of Interaction:	learning to selecting
Mode of Retrieval	recognition to specification
Resource Considered:	inforamtion to meta-information, [BMC93, p. 325]

Considering each of these as a dichotomous variable, Belkin, Marchetti and Cool [BMC93] and Belkin, Cool, Stein, and Thiel [BCS+95] consider 16 information-seeking strategies (ISS). In their model, the ISS used by a person may change as the information seeker's situation changes. Belkin et al. [BCS+95] describe an example of this:

> "For example, a person may not be able to specify the title or author of a book she is looking for, but may remember its approximate shelf location. In order to recognize the item, this person might go to this location and scan the shelves (ISS5). While at the shelves, she may find some other book or item that seems relveant. This might prompt her to search for other similar items, but first she needs to know how to find these items. In order to learn how other items have been characterized in the information system, she may turn to a meta-information resource, such as a catalogue, to identify the subject terms under which the item has been indexed (ISS12). Having finished this step, this user would be able to continue the information-seeking episode by specifying these terms in a search for other relevant items (ISS15)." [BCS+95, p.381]

---

[2] Reprinted from Information Processing and Management, Vol. 35, Vakkari, P., Task complexity, problem structure and information actions: Integrating studies on information seeking and retrieval, 819-837, Copyright (1999), with permission from Elsevier.

Belkin et al. [BCS+95] further explore how this characterization of information seeking can be used in an interactive system based on the concept of the user and system engaging in a dialogue.

2.5.3   Web Information Seeking Behavioral Model – Choo, Detlor, and Turnbull

Choo, Detlor, and Turnbull [CDT00] conducted a review and empirical investigation of how modes of scanning/browsing as defined by Aguilar and expanded by Weick and Daft (undirected viewing, conditioned viewing, informal search, and formal search) are realized through the set of information seeking behaviors outlined by Ellis (starting, chaining, browsing, differentiating, monitoring, and extracting). In the study by Choo et al., the Web behaviors of 34 business workers were collected and analyzed during two-week periods of time. Each participant completed a questionnaire, had data collected through a custom software package that logged URLs and browser actions, and completed an interview about critical incidents of Web information seeking. Sixty-one information-seeking episodes were selected and categorized according to the modes of scanning and the information seeking behaviors used. Choo et al. [CDT00] reported the most frequent behaviors for each mode:

- Undirected viewing – starting, chaining
- Conditioned viewing – differentiating, browsing, and monitoring
- Informal search – differentiating and extracting
- Formal search – extracting

These results are interesting because they illustrate a research that has combined several models from information science and empirically investigated their application to Web information seeking.

2.6   Studies of Information Refinding and Reuse

Few studies have specifically examined how people refind information on the Web and how the process and strategies used in refinding relate to those used when initially finding information. In this section, several studies of refinding and related topics are reviewed.

2.6.1   Studies of Refinding

2.6.1.1   *Maglio and Barrett, 1997*

The most relevant prior study was an examination of how people search the Web by Maglio and Barrett [MB97a] in which participants performed identical search tasks on two different days. On the second day, before each task, participants verbally recalled their search actions from the first day. This study had a small sample size (7 participants and 3 searches each day), but was conducted as a controlled study. Two main results were reported. First, that people rely on routines for searching, even to the point of remembering search steps on the second day that were not actually performed on the first day. The second result was that while participants could not always reconstruct a search path exactly, that certain "key" nodes, or *waypoints*, were remembered and play an important role in refinding.

2.6.1.2   *Capra and Pérez-Quiñones, 2002*

We confirmed and extended the results of Maglio and Barrett [MB97a] in a study of collaborative information refinding [CP05a] [CP03]. The study involved pairs of participants conversing over the telephone. One participant (the *retriever*) helped the other participant (the *user*) refind Web information that the user had previously found during a session held one week earlier. To facilitate this, the retriever sat at the computer that the user had used when finding the information. The retriever had access to a complete history of the Web browsing from the user's finding session. In addition, during the initial finding session, the user was able to make and categorize persistent annotations similar to digital post-it

notes on the Web pages. These annotations and their categories were also available to the retriever during the refinding session. The study was conducted with six pairs of participants. Each of the six pairs completed four tasks.

This study design allowed us to gain insight into the strategies used to refind and present information, and into how domain artifacts and contextual information were used to help the refinding process. The conversations were transcribed and coded for: 1) spoken references to waypoints and annotations, 2) conversational grounding events, and 3) specific requests for information.

Both users and retrievers used waypoints extensively to help guide refinding and to establish common ground. Twenty of 26 conversations included waypoints (76.9%). Additionally, we identified three types of waypoints (page titles, URLs, and page descriptions) and characterzied how the users and retrievers used them in refinding.

2.6.1.3 *Capra, Pinney, and Pérez-Quiñones, 2004*

Additional work on refinding has been conducted by our research group at Virginia Tech. Mary Pinney [CPP04] conducted a controlled-study to examine aspects of refinding, especially with regard to post-valued recall (PVR). She did not observe any difference in refinding between participants who were and were not specifically informed that they would need to refind information later. However, motivating such a need is difficult in controlled conditions and so these results may not be conclusive. The results of the study did indicate differences between task types – participants had less accuracy refinding information for exploratory tasks than for tasks to refind particular Web pages [CPP04].

2.6.1.4 *Teevan, Alvarado, Ackerman and Karger, 2004*

Teevan, Alvarado, Ackerman and Karger [TAA+04] conducted a study in which they interrupted participants several times a day over the course of a week and inquired about their recent information seeking for files, email messages, and Web pages. They observed instances of participants *teleporting* directly to information and also of *orienteering*, or making use of cues in the environment to relocate an item. They also reported that participants used keyword search in only 39% of the information seeking episodes they examined [TAA+04].

2.6.1.5 *Boardman and Sasse*

Boardman and Sasse [BS04] examined the personal information management practices of 31 participants in a study that examined data and behaviors across tools for files, email, and Web. Among their many observations and results, they found that there was some overlap of strategies and structures between tools (i.e. creation of similar folders), but that "most folders did not overlap" [BS04, p. 589]. They also note that participants used multiple PIM strategies both within collections of data and across tools [BS04, p.589].

2.6.1.6 *Aula, Jhaveri, and Käki*

Aula, Jhaveri, and Käki [AJK05] examined the search and re-access strategies of 236 experienced Web users. They identified a number of "key strategies" [AJK05, p. 587] that these users employed in their searching (e.g. the use of multiple windows/tabs, and a recognition of the benefits of categorized information). They also discuss several important strategies they observed for information re-access (e.g. the use of search engines – although participants noted problems with this approach, and the use of bookmarks). Alua et al. [AJK05] also note several issues with existing tools and strategies (e.g. problems with history tools, and the difficulties that users have understanding aspects of how Web search engines work).

2.6.2   Web Page Revisitation

2.6.2.1   *Revisitation and Personal Information Retrieval*

Several efforts have examined patterns of Web page revisitation and found that users have a set of pages that they tend to revisit often [TG97] [CGJ+03]. Several companies are investigating and deploying tools for refinding personal information such as documents, email messages, and Web pages. Among these are Microsoft (Stuff I've Seen project) [DCC+03], Google (Google's desktop search program), and A9 (A9 toolbar).

Several research projects and studies have explored how to help people recognize and revisit Web pages that they have seen before. Kaasten, Greenberg, and Edwards [KGE02] conducted empirical studies as the basis for guidelines for the recognition of previously seen Web sites and Web pages. One set of guidelines presents the sizes of thumbnails needed to achieve various recognition rates. For example, a 60% recognition rate of a Web site can be achieved given a thumbnail image that is 96x96 pixels. Similarly, they developed guidelines for the number and positions of characters needed to achieve various recognition rates of textual URLs and Web page titles. Woodruff, Rosenholtz, Morrison, Faulring, and Pirolli [WRM+02] examined the use of text summaries, thumbnail images, and enhanced thumbnails in information finding tasks. The enhanced thumbnails had certain keywords made readable. They found that enhanced thumbnails provided the best overall performance, but that for some tasks text summaries were easier to use than thumbnails and vice-versa for other tasks [WRM+02]. Two separate projects – TopicShop [ATH+03] and DataMountain [RCL+98] have explored interfaces that include the use of thumbnail images in organizing and understanding Web pages.

Cockburn, Greenberg, Jones, McKenzie, and Moyle [CGJ+03] discuss issues with Web page revisitation based on a study examining the Web activities of 17 users during a 119 day period [CGJ+03, p. 163]. Based on their data, they report, "more than 80% (on average) of Web page visits being to pages previously visited by the user" [CGJ+03, p.180]. They also describe alternatives to the standard history mechanisms such as "providing a complete, temporally ordered list of previously visited pages." [CGJ+03, p.180]

2.6.2.2   *The Keeping Decision*

William Jones describes "the keeping decision" as whether to keep some found information or not, and asserts that, "the keeping decision is fundamental to the management of personal information and fundamentally difficult to do." [Jon04, Conclusion section]. He examines the keeping decision as a signal detection theory problem, weighing the costs of keeping (e.g. storage space) and the costs of keeping mistakes (e.g. too much information, not the right information). In an example of how the costs of keeping mistakes may be decreasing, he notes that even if people take no specific steps to keep information found on the Web, "they can often easily return to this information…" [Jon04, Conclusion section]. However, he also notes the limitations and costs of having to remember to return to information and of having kept too much information.

2.6.2.3   *Revisitation and Personal Information Retrieval*

Tauscher and Greenberg [TG97] conducted an extensive and early study of how people revisit Web pages. They studied 6 weeks of Web browsing data from 23 users. Among their findings were that Web pages are frequently revisited (revisitation rate of 58%), but that only a few Web pages make up most of the revisited pages. They also found that many pages are only visited once (60%) or twice (19%) [TG97, p. 131]. Of particular interest for refinding research, they observed that some revisits were to Web pages that had not been visited recently.

2.6.3   Finding and Reminding

2.6.3.1   Finding and Reminding – Barreau and Nardi; Finding and Reminding Reconsidered – Fertig, Freeman, and Gelernter

Barreau and Nardi [BN95] reported on two interview studies conducted in 1993 and 1994 that examined how users "organize and find files on their computers." [BN95, p. 39] The first study examined the file storage behaviors of seven managers that used mostly DOS-style filesystems. The second study reported on fifteen Macintosh users in a variety of job positions. In both studies, similar results were found: users preferred a style of location-based search as opposed to using a keyword-based search utility, file location on the desktop served as a reminding function, and three types of information were observed: ephemeral, working, and archived. In addition, users did not seem to stick to elaborate filing schemes to archive files.

Location-based search as described by Barreau and Nardi is similar to the "orienteering" behavior observed by Teevan et al. [TAA+04] and is related to Lucy Suchman's notion of situated action [Suc87]. Barreau and Nardi observed, 'the user takes a guess at the directory/folder or diskette where she thinks a file might be located, goes to that location, and then browses the list of files or array of icons in the location till she finds the file she's looking for. The process is iterated as needed." [BN95, p. 40]

Barreau and Nardi note that because of the reminding function of document placement, a system that does not include the notion of location (i.e. one that relies purely keyword search) will lack the important reminding function that location adds. To elaborate on their point, in effect, location is serving as an important type of context on which users often rely to help refind files. Since users often rely on this context when refinding, systems that do not include the context might not be used. Although Barreau and Nardi's research was done over 10 years ago, the results are especially relevant today. New filesystems and refinding tools that make keyword searching of personal documents easy need to find ways to incorporate context and location into the refinding process. One example is the notion of having a location, but that the location is a "query" of metadata. For example, Google's Gmail (gmail.google.com) allows users to add labels to messages [Pog04].

Barreau and Nardi also note three types of information: ephemeral, working, and archived. Ephemeral information is information that is needed for a short length of time – email, lists, notes, memos, etc. Working information is information that has recognized importance and is important to the user's current work tasks. Archived information is information that is infrequently accessed and may include final versions of documents and reports.

Fertig, Freeman, and Gelenter [FFG96] raised several issues in a response to Barreau and Nardi's findings. One of the main points they questioned was whether Barreau and Nardi observed a predominance of location-based finding because of a real preference for this method, or because of a lack of breadth in the studies and a lack of good tools to provide keyword search at the time the study was conducted. This is a good point that is still relevant in current research on finding and refinding. Teevan et al. [TAA+04] reported that only 39% of the refinding instances they examined in a study of 15 users refinding behaviors for Web, email, and files used keyword-based search. However, of the three domains examined, only pages on the Web have robust search tools (e.g. Web search engines). Windows does provide a search tool for files, but it is limited to searching filenames, dates, and times. Email clients often provide search tools, but they too are often limited to searching email headers such as subject lines. Recently, more sophisticated "personal search" tools such as Google Desktop and Apple Spotlight are being released. Many of these tools index all documents, email, and saved Web pages stored on a user's hard disk and allow users to issue keyword queries over their personal collection of documents in a similar way that they use Web search engines. The deployment and adaptation of these tools will provide users keyword search capabilities for files and email that rival those for Web pages. The effects of this on finding and refinding behaviors is not yet understood and should be a topic of future research. Fertig et al. also point out that using file locations as reminders is somewhat "ad hoc" and is constrained by a lack of labels and screen space. They point to research on dynamic queries and virtual directories as innovative ways to explore file spaces that need more exploration. These concepts (dynamic queries and virtual directories) appear to have influenced new designs for filesystems such as Microsoft's WinFS [Riz04] and the organization of email in

Google's Gmail (gmail.google.com) [Pog04]. As tools such as these are deployed and used, researchers should explore how they affect information seeking behavior.

2.6.3.2   *Ravasio, Schär, and Krueger, 2004*

Ravasio et al. [RSK04] followed up on the work of Barreau and Nardi, Whitaker and Hirshberg and others to examine the behaviors and problems that users have with desktops, file systems, and information finding and refinding. They interviewed 16 computer users (12 Windows and 4 Macintosh) about their working environments. They found that users viewed archives and maintenance of their files and file systems to be important tasks and that users relied on their archives. They found that classification and labeling were still tasks for which users, "expended considerable cognitive effort" [RSK04, p. 164]. They found that skilled users made use of desktop screen space to group icons by proximity based on themes and that the desktop was often used for temporary storage and verified the continued existence of ephemeral, working and archived information. Users with lower skills did not use the desktop as extensively and may not even realize that documents can be placed there. Users expressed a need for a usable and useful way to add notes and comments to documents and felt that existing software solutions were not sufficient. Ravasio et al. [RSK04] also observed a prevalence of manual search (similar to the findings of Barreau and Nardi), the use of spatial structures for organization (similar to Malone), and the "reluctant use of search tools" [RSK04, p. 170]. Users were comfortable using Web search tools, but reported that when searching their personal space, they preferred manual search because, 1) internet search was used to look for new information, while searching their personal space was to refind a known item, 2) manual search helped provide "an up-to-date overview and orientation of the contents' structure" [RSK04, p. 171], and 3) there are limitations with the interfaces for current built-in OS filesystem search tools such as the one in WindowsXP [RSK04, p.170, p.175].

2.6.3.3   *Malone, 1983*

Malone [Mal83] conducted one of the first studies of desktop file organization aimed at informing the design of computer filesystems and tools. In his study, published in 1983, he examined how people organized information on their physical desktops and office surfaces such as tables and cabinets. He interviewed 10 people that performed a large amount of work at their desks. He codified messy and neat offices and identified two forms of organization: files and piles. He observed that in addition to the finding function that the spatial organization of items on a desk provides, there is an important reminding function on which users rely. He noted that people have trouble figuring out how to classify and label items and that this can be a barrier to filing. He hypothesized that this difficulty may be one reason that people have many loosely classified piles of paper on their desks [Mal83, p. 108].

Malone suggested several ways that computers could help people in file organization and management based on studies of physical desktop use. Among these were: support for multiple classification of a single document, allowing deferred classification of a document by storing it on a desktop as users currently do, the automatic classification of documents based on metadata such as email fields, file access times, etc., and allowing document retrieval and searching based on multiple fields.

2.6.3.4   *Whittaker and Hirschberg, 2001*

Whittaker and Hirschberg [WH01] looked at how employees in an industrial research environment managed archives of paper information. Their study coincided with the employees moving to a different office space, so they were able to inquire about the decisions made about whether to keep paper information and the quantities of information both before and after the move. Fifty employees completed an on-line survey and 14 participated in follow-up interviews. They found that despite increasing availability and use of electronic information storage and publicly available electronic information repositories, that people kept and valued large amounts of paper information. Workers kept 78% of their paper archives and spent an average of almost nine hours reviewing their data for the move. They also found that people retained large amounts of unprocessed or under-processed data due to information overload. Participants reported having trouble evaluating the value of information, and would save it for deferred evaluation in case it was needed. However, cleaning out workspaces was not often done spontaneously – 74% had not performed a clean-out in the previous year and 84% of clean-outs were due to

job changes or office moves. People did keep some paper information because it was unique, such as working notes, archives of completed projects, and legal documents. However, 15% was unread and 36% were copies of publicly available documents. These copies were kept for reasons of availability, reminding, and in some cases, concerns about relying on publicly available sources for important information.

Whittaker and Hirschberg also investigated the practices of filers and pilers. They note that many of the participants in their study kept both files and piles, so they separated them based on a 40% threshold. Surprisingly, they found that filers had larger paper archives than pilers. They theorized that this may be due to filers prematurely filing information before knowing its long-term value. Then, once filed, it remains filed. Another interesting finding was that pilers tended to access their archives more frequently, perhaps due to availability.

Whittaker and Hirschberg conclude that there are three major problems that users face when dealing with information, in paper or in other forms: "(a) determining the value of incoming information, (b) deciding whether and how to categorize that data, and (c) in electing whether to keep incoming information available in their immediate workspace or file it away for later access". This last point is related to "the keeping decision" and, as Whittaker and Hirschberg note, is complicated by information overload and the changing value of information (related to Wen's notion of post-value recall [Wen03]).

2.6.3.5   *Graphics, Visualization, and Usability Center, 1998*

A 1998 study [GVU98] conducted by the Graphics, Visualization, and Usability Center (GVU) at Georgia Tech found that 73.6% of the respondents included "gathering information for personal needs" among their responses to the question, "What do you primarily use the Web for?" (GVU Center at Georgia Tech 1998, "Primary Uses of the Web" section). This question was a multiple-choice, "check all that apply" question in the "Primary Uses of the Web" section of the GVU survey [GVU98]. Other possible responses to the question included: Education (61.4%), Shopping/gathering product information (52.4%), Entertainment (60.1%), Work/Business (65.9%), Communication with others (not including email) (35.4%), Wasting time (37.0%), and Other (8.7%). "Gathering information for personal needs" was the highest response with 2421 out of 3291 participants (73.6%) including it in their marked choices [GVU98].

In the same survey, 45.4% of the respondents included, "Not being able to find the information I am looking for" in their responses to the question, "What do you find to be the biggest problems in using the Web?" [GVU98, "Problems Using the Web" section]. This question was a multiple choice, "check all that apply" question in the "Problems Using the Web" section of the GVU survey. There were 19 possible responses to this question. "Not being able to find the information I am looking for" (1493 out of 3291 participants, 45.4%) was one of nine responses with higher than 40% of the participants including it in their marked choices [GVU98].

Also in the GVU survey, three additional responses to the question, "What do you find to be the biggest problems in using the Web?" are of interest:

> 30.0%   Not being able to find a page I know is out there
> 16.6%   Not being able to return to a page I once visited
> 27.6%   Not being able to efficiently organize the information I gather
> [GVU98, "Problems Using the Web" section]

These results suggest that re-finding once-found Web pages is a problem and that users have trouble organizing information found on the Web.

2.6.4   Bookmark Usage

Jones, Bruce and Dumais [JBD01] observed[3] that users in their study did not make extensive use of bookmarks. However, earlier research with larger sample sizes has shown that users do create bookmarks and has investigated their use and organization. In the 1998 GVU Web Survey, only 2.2%[4] of users reported not using bookmarks or favorites and 57.7% reported having more than 50 items in their bookmark lists [GVU98]. Abrams, Baecker and Chignell [ABC98] reported in 1998 on an extensive study of bookmark usage. Among their findings was that although bookmarks are relatively easy to create, many users don't organize bookmarks (approximately 37% of their respondents said they did not organize their bookmarks) [ABC98, p.44]. The 1998 GVU Web survey provides a slightly different view of bookmark organization: 75.5% of respondents indicated that they had manually re-arranged bookmark entries, 76.9% reported that they had created folders of related items, 65.6% indicated that they had changed an entry's title, and 82.1% reported having deleted an entry [GVU98, "Favorites/Bookmarks Usage Patterns" section]. In the same survey, on a question regarding favorites/bookmark problems, 19.9% of the respondents cited "there is no way to quickly store a new page in the place I want it" and 20.2% cited "there is no good way to organize the pages" as problems with the favorites/bookmark utility in their browser [GVU98, "Problems Using Favorites/Bookmarks" section]. These results have implications for the types of engagement that can be reasonably expected from users to enter explicit context to help with information refinding. As Abrams et al. note, "Organizing bookmarks is labor intensive, requires extensive time, and is difficult to do" [ABC98, p.44]. A recommendation from this observation is that tools are needed that "minimize the effort needed to organize bookmarks… Systems should provide users with an immediate filing mechanism so that bookmarks can be filed at creation-time" [ABC98, p.47]. An additional recommendation is that "an integrated bookmark management system should fit naturally into the Web browser" [ABC98, p.47].

2.7   Tools for Finding and Organization

A number of tools have been developed and studied to help users find and organize information on the Web. In this section, Web browsing tools and assistants related to information finding, refinding, and PIM will be reviewed.

A handful of projects have examined the recognition of "nuggets"[5] [LNW99], or pieces, of information embedded in documents and techniques to allow users to perform context-sensitive tasks using the recognized information nuggets. Several of these projects were investigated in the context of collaborative agents and have explored techniques for unobtrusively displaying information and allowing users to manipulate it. Researchers at Apple Computer explored this area with their Apple Data Detectors [NMW98] and the related LiveDoc [MB98] and DropZones [BM98] projects. Pandit and Kalbag [PK97] at Intel Labs developed a similar system called the Selection Recognition Agent. These systems provide users an ability to perform a set of customized, relevant actions on pieces of information that are recognized by the system. For example, if a document contained a phone number, the system might parse the text, recognize the phone number, and provide options such as "call this person at this number" [BM98, p.60], or "add this person to my address book" [BM98, p.60], possibly through a contextual menu obtained by right-clicking on the email address in the text. The use of contextual menus allows unobtrusive display of contextual information and allows users to manipulate information without interfering with other tasks[6]. These systems often include an interesting use of implicit and explicit input; typically, the information

---

[3] The authors note that they make this point about low bookmark usage "tentatively" [JBD01, p.123] because of the way their study was conducted.
4 In response to the question, "Approximately how many items does your Favorites/Bookmarks contain?", 2.2% responded "I don't use Favorites/Bookmarks" [GVU98, "Number of Items in Favorites/Bookmark" section].
[5] Lieberman has used the term "nuggets" to describe, "semantically meaningful and syntactically recognizable items embedded in a larger body of unstructured information" [LNW99, p.116]. In his description, semi-structured information is information that contains such "nuggets". I will use the phrases "information nuggets" and "information pieces" interchangeably in this section to refer to the same types of "nuggets" that Lieberman defined.
[6] Not all the systems described here use contextual menus, but they are a prominent feature in several of the systems.

nuggets are implicitly recognized by the system, but actions are not taken by the system until the user provides explicit input.

2.7.1   Intel Selection Recognition Agent

The Intel project [PK97], called the Selection Recognition Agent (SRA), intercepts operating system events to determine when text has been selected in a window. It passes selected text along to a recognition process that looks for information nuggets such as "geographical names, dates, email addresses, Usenet newsgroup name components, world-wide Web site names (URLs), and phone numbers" [PK97, p.47]. The SRA can also acquire text when it is copied to the clipboard. Then, by right-clicking, the user can bring up a contextual menu of actions that can be taken on the information nugget such as opening a calendar program to the date selected, or starting an email program with a message addressed to the email address selected. The SRA is application-independent and works with Microsoft Windows 95. Timesavings and reduced keystrokes/mouse-clicks to accomplish tasks are cited as benefits of the SRA [PK97].

2.7.2   Apple Data Detectors

Apple Data Detectors (ADD) [NMW98] are very similar in nature to the Intel Selection Recognition Agent. ADD allows users to select regions of text in documents and perform sets of actions involving information nuggets that are found in the selected text. Unobtrusiveness was a major design goal of ADD, so much its operation is conducted "behind the scenes" and it is "invisible until needed" [NMW98, pp.98-99]. Contextual menus are used to display a list of recognized information structures found in the selected text and the actions that can be taken on them [MB98]. Apple Data Detectors can be extended by writing new detectors and action scripts. A scripting language can be used to manipulate data and to interface directly with any existing "scriptable application" [NMW98, p.101]. At the time of the article (1998), writing new detectors and action scripts was a job for programmers, but the researchers were investigating the development of end-user programming tools. The Grammex, "Grammars by Example" [LNW99, p.116], tool was an apparent result of this effort. Grammex allows users to create context-free grammar rules by example. Such grammar rules are important components for being able to customize systems such as Apple Data Detectors that parse text and use grammar rules to identify information of interest.

2.7.3   LiveDoc and DropZones

Apple's Advanced Technology Group worked on two projects, LiveDoc and DropZones that were related to Apple Data Detectors. These projects were based around the idea of "structure detection" in documents [MB98, p.53]. LiveDoc [MB98] was a system developed to extend ADD and address some of its limitations. Like ADD, LiveDoc operates in the background to find information nuggets, but it does not require users to select the text to be analyzed as ADD does. Instead, LiveDoc parses the entire visible portion of a document and will highlight all recognized structures if the user presses a function key. This automatic structure detection and highlighting saves users from having to select a section of text and wait for it to be analyzed. Performing structure detection as a background process can allow more sophisticated processing to take place since the system does not need to complete its parsing in the "very short period of time in which users are willing to wait for a pop-up menu to appear" [MB98, p.55]. Highlighting the recognized structures allows them to be treated as direct manipulation objects [MB98]. LiveDoc was implemented using a client-server architecture that included application clients, a LiveDoc manager, and an Analyzer Server.

DropZones [BM98] was a project that explored issues beyond LiveDoc, especially in the area of enabling users to combine information nuggets to perform more advanced tasks and manage relationships between items. DropZones employs the use of "assistants" [BM98, p.60] that are able to process and combine particular pieces of information (found by LiveDoc) within a context. For example, the Email assistant can accept names and email addresses. The user can select and then drag information pieces over icons for a

variety of assistants. If the assistant can process the information, it highlights itself. If the user then drops the information on the assistant, it will display a set of actions that can be taken in an "Assistant window" [BM98, p.61]. A main feature of Drop Zone assistants is their ability to compose (combine) pieces of information. For example, the "Meeting Assistant" (shown in figures as the "Appointment Assistant") [BM98, p.62] can compose a date, time, and place into an appointment which can be added to the user's calendar. DropZones includes mechanisms to represent semantic information about the information found by LiveDoc. It uses a "frame-based object system, augmented by relational axioms to express facts about these objects" [BM98, p.61]. Facts and axioms are used by the assistants to help process information objects and draw inferences. For example, by using an axiom, the email assistant can determine an email address if given a phone number. The authors note that this aspect of DropZones runs into traditional problems of knowledge representation, but that it is offset somewhat by the limited scope of the information being processed [BM98].

Context shared by the user and the system can assist users in refinding information when it is needed. Making context explicit can help the user describe their information needs and can help the system focus and describe information it has stored. Drop Zones provides assistants to help users explicitly organize and combine pieces of information. These capabilities are supported to some degree in Web annotation tools that integrate with Web browsers. However, the assistants-plus-direct manipulation approach taken by DropZones is an interesting interface for entering explicit context.

2.7.4   Cyberdesk

Cyberdesk is a framework for integrating information across applications that makes use of contextual information [DAW98]. Cyberdesk uses the user's context to present a set of possible actions that the user might wish to take in desktop applications. For example, if the name of a person were selected in a Web browser, Cyberdesk might provide options such as, "search for the selected text using AltaVista, search for a phone number and mailing address using Switchboard, lookup the name in the contact manager" [DAW98, p.48]. New applications can be integrated into Cyberdesk by writing wrapper functions that register services provided by the application and help Cyberdesk execute application functionality. Cyberdesk has chaining and combining capabilities similar to those found in DropZones. Chaining is a process by which Cyberdesk can use one piece of information to find another. For example, if a user selected a name, this name can be looked up in a database to provide an email address. Combining can be used to enable more robust actions that involve multiple information pieces. For example, a name and address could be combined and used to lookup a phone number.

2.7.5   Observations about Tools for Finding and Organization

The projects reviewed in this section (2.7) all explored ways to allow users to take actions on recognizable, semi-structured, information nuggets. The use of contextual menus, function keys, and direct manipulation to take actions involving the information nuggets is a type of explicit input by users regarding the information. This input could be used to develop explicit user-system shared context. The information nuggets identified by the system can also be viewed in terms of a user's need to store and re-find semi-structured information. Some of the actions and functions supported by the projects reviewed here include mechanisms for storing information for later re-use – saving an appointment in a calendar for example. However, many information nuggets don't fall into pre-defined application categories such as email, calendar appointments, and address books and could benefit from a general-purpose mechanism to store and retrieve information nuggets.

The projects reviewed in this section (2.7) provide insights and contributions about how to create easy-to-use, unobtrusive interfaces to allow users to identify and manipulate information nuggets. However, additional mechanisms for entering explicit context may also be needed to support robust user-system shared context. Web annotation systems also contribute ideas about how to develop good interfaces to allow users to explicitly identify and label information for later re-access.

2.8   Personal Information Spaces and Automatic Indexing Tools

The previous section focused mainly on systems that recognize information nuggets in documents and allow users to perform actions on this information (mostly at the time it is being viewed). A number of other projects have focused on helping users automatically archive and index information they have viewed, with a goal of making it easier to re-find and retrieve at a later date. A number of these projects have been Web-oriented, but several have attempted to index information from a wide variety of sources.

2.8.1   Remembrance Agent and Margin Notes

Remembrance Agent [RS96] and Margin Notes [Rho00] are related projects developed by Brad Rhodes at the MIT Media Lab. The Remembrance Agent (RA) is a system that works in conjunction with Emacs and observes the content of the document the user is currently working on [RS96]. Based on this, the RA provides suggestions of other documents in the user's collection that are possibly relevant to the current section of the text. To make these suggestions, it maintains an index of all the documents, emails, text files, etc. that the user views. Suggestions are made in a separate Emacs window. Measuring the utility or usefulness of the suggestions is difficult, especially because "there is a large difference between relevant suggestions (those that have a strong relation to the user's current context), and useful suggestions" [RS96, "Evaluation" section, para. 1]. They comment that the best way to evaluate such systems is in the "context of actual use" [RS96, "Evaluation" section, para. 1].

Margin Notes is a related system that is integrated into Web browsing rather than into Emacs [Rho00]. Margin Notes indexes "email archives, notes files, and other text files" [Rho00, p.220] and uses this index to make suggestions of documents that might be relevant to the Web page that a user is currently viewing. Margin Notes is implemented as a proxy that dynamically annotates Web pages by displaying information about and hyperlinks to potentially relevant documents in a margin that it adds to each Web page displayed [Rho00].

Both Margin Notes and the Remembrance Agent are presented as tools to help users recall potentially relevant documents from their personal store of information. They attempt to anticipate a user's information needs and present relevant information.

2.8.2   Haystack

Haystack is a system for gathering, storing, and indexing a personal store of information [AKS99]. Haystack automatically performs a number of these functions, yet also tries to expose and exploit relationships and connections that it can discover among the data stored. A main idea behind Haystack is to provide an information storage and retrieval system that is customized and adapts to a particular user.

Haystack implements a very generalized data model that helps it annotate information and make connections between various pieces of information. The generalized nature of the data model allows Haystack to store a wide variety of information types. Information is stored in nodes of a graph (the nodes are called straws) [AKS99]. Straws are a superclass of an inheritance hierarchy that includes needles, bales, and ties as subclasses of straws. Needles store "a single primitive piece of information" [AKS99, p.416] such as the text of a document or the document type. Bales are used to group together other nodes, such as to group all the information related to a particular document. Ties are used to connect other straws. Ties are really the edges of the graph, but since they too are straws, they can have things tied to them.

Haystack uses several types of "observers" [AKS99, p.418] to help it make inferences and connections among data without requiring explicit user input (although Haystack also includes mechanisms for users to make annotations and connections by hand). Web and email proxies are used to observe Web pages and email documents viewed by the user and to add them to the user's Haystack. As documents are added, Haystack can annotate them with a "'visited' time" [AKS99, p.420] to allow users to make queries based on the time a document was viewed.

Queries in Haystack are treated as first-class objects. Query observers watch a user's queries to their Haystack and store these as Haystack objects that can be annotated or connected to other objects. Haystack can use these stored query objects to help refine query results, reinforce relevant documents, and provide results to other queries.

Data are added to a user's Haystack not only through the addition of documents that they view, but also through a number of "data driven clients" [AKS, p.419] that operate on "information that already exists in the repository" [AKS, p.419] to add new information. Examples of data driven clients include: fetch clients that can retrieve related documents and information, such as fetching a URL that is mentioned in a document; type inference clients that guess the type of a document and add this information, such as a PDF file or HTML Web page; extractor clients that "attempt to extract textual information from the retrieved documents" [AKS99, p.419] such as converting postscript into text; and field finder clients that "extract various pieces of meta-data" such as "the to, from, and subject lines in an email message" [AKS99, p.419].

Human annotation is another source of information that can be added to a Haystack. Several interfaces are supported to allow users to add annotations to information stored in their Haystack: "a custom (personal) Web server, a Java based GUI, a command line tool, and an emacs interface" [AKS99, p.421]. The authors note challenges in getting users to devote effort to annotation: "This variety of access mechanisms will perhaps encourage the integration of annotation into the user's workflow. However, significant user studies will be required to refine this approach" [AKS99, p.421].

2.8.3   Watson

Watson is a system that observes users interactions with desktop applications such as a word processor and Web browser and uses this context to try to make pro-active recommendations of other information relevant to the user's current context and activities [BH00]. The text of open documents is processed and used to form queries that are sent to a variety of Internet information sources. The results of these queries are processed and presented to the user through Watson's user interface window. Waston is designed around the idea of using activity on the desktop as implicit input (context) for anticipating and assisting with the information retrieval needs of the user. Users can also input explicit queries using Watson. Watson can complement these explicit queries with contextual information from the user's open documents in an attempt to form more context-relevant queries (and results).

2.8.4   The Information Programming Toolkit

The Information Programming Toolkit (IPtk) is a project at IBM [FBC+02]. Describing the project, the authors note, "To support how people naturally and routinely organize information, computers ought to be able to reflect the categories, relationships, and cues that people rely on when thinking about and remembering facts" [FBC+02, p.190]. The IPtk appears to be focused on the programming architecture to support information use. The IPtk has been used to develop a "personal library" [FBC+02, p.191] application that uses intermediaries and IPtk "promiscuous indexing" [FBC+02, p.190] to create an index of information that a user has viewed from a variety of different sources. The IPtk appears to incorporate context into its indexing; the authors describe queries based on user recall such as "the note I took while talking to Barney on the phone" [FBC+02, p.191], or an email message "received from Fred before my trip to Acme," [FBC+02, p.191] although it is unclear how much of this functionality is supported in their current implementation.

2.8.5   Vistabar

Vistabar is a Web browsing assistant that provides functionality to save and index Web pages, to allow users to search over their personal repository of viewed Web pages, to find Web pages related to a particular page, and to label and view classification and annotation information for pages [MB97b]. Vistabar works in conjunction with Netscape and Internet Explorer by using published APIs to "monitor and control" the browser [MB97b, p.131]. Vistabar attaches a small button bar at the bottom of the regular

browser window to give users access to its features. Vistabar features are summarized and discussed below.

Saving and indexing viewed Web pages – Vistabar can save and index all the Web pages a user views. Two alternative modes are provided if the user does not wish to index all pages. Explicit mode can be used to save the current page by pressing the "Remember" button on the Vistabar [MB97b, p.132]. Temporal mode can be configured to save all pages that were viewed for a minimum length of time [MB97b, p.132].

Searching saved pages – Vistabar allows users to search their personal store of saved pages. Vistabar indexes the stored pages and provides query capabilities. [MB97b, pp.132-133].

Finding related pages/categories – A "Similar" button is provided that invokes a simple scheme that extracts the most common words from the page and uses them as a search query to AltaVista. A more advanced technique is used to support a "Classify" button that attempts to find related categories in the Yahoo! hierarchy [MB97b, p.133].

Classification and Annotation – Three buttons are provided to support classifications and annotations [MB97b, p.136]. First, an "Add Comment" button allows users to fill out a form with a comment about the Web page and to specify a category for the page. A list of predefined categories is provided, but users can also create their own. Second, a "Comments" button can be used to view comments associated with a page. The button lights up if the page contains a comment. Third, an "Auto Bookmark" button is provided that will cause Vistabar to make a guess at an appropriate category for the current page based on "the profiles of the categories currently in the system" [MB97b, p.136].

2.9   Web Annotations

Web annotation tools allow users to make various types of annotations on Web pages. These systems provide a method for users to make explicit notations and to flag important information. In addition, they tend to be well integrated with existing Web browsers and tools. Several commercial products exist that provide users an ability to highlight text, draw on pages, add free-text comments, classify, and save Web pages they have viewed. Some of these systems are aimed at individual users, while others focus more on collaborative uses of Web annotations. In this section, several Web annotation systems of interest will be described.

Web annotations have been around since at least 1993 [NCS93] [Mos98] [KKP+01] and many different annotation systems have been built and explored. Researchers have explored the use of Web annotations for information retrieval [DV00], queries [GPS99], personal information organization [DV01], and collaboration [CGG00]. Two good surveys of Web annotation systems are presented in [HLO99] and [KKP+01]. There is also a good list of links to annotation systems on the CritSuite Web site (http://crit.org) in the "Related Projects and Software" section.

2.9.1   CritSuite

CritSuite is a set of tools including CritLink that allow users to leave publicly viewable annotations about any Web site [Yee02a] [Yee02b] [Yee98]. Web pages are viewed through a mediator (the crit.org Web site can act as a mediator). The mediator fetches pages that users request, looks up annotations for the page in its database, combines the annotations with the page requested, makes a few additional modifications to the Web page to facilitate seamless continued use of the mediator, and returns this combined page to the user [Yee98]. An interesting interface feature of CritSuite is its use of color markers to surround text that has been annotated.

2.9.2   Annotea

Annotea is an open system being developed through the World Wide Web Consortium (W3C) to support both personal and shared Web annotations. Motivated by the semantic Web initiatives [W3C02], Annotea is part of a goal to build "the infrastructure needed to handle and associate metadata with content" [KKP+01, p.623]. Annotea uses Web technologies including RDF (Resource Description Framework), XPointer, and XLink to represent, access, and display user annotations to Web pages. Annotea uses a client-server architecture and incorporates the notion of local and shared annotations. Annotea authors have been using the W3C Amaya Web browser to explore the client side interface to annotations. Annotations can be associated with an entire page, a selected section of a page, or a specific cursor point. A pop-up dialog box allows users to enter text for the annotation and also to view the annotation's metadata (such as creation date and time, author, last modified, annotation type, etc.). Amaya displays annotations by embedding a small icon of a pencil at the point in the document where the author placed the annotation. Annotations are stored independently from the Web pages they annotate. Local annotations are not shared with other users and are stored locally on the user's hard drive. Shared annotations are stored on annotation servers and may be accessible by other users. Thus, shared annotations can be used for collaborative activities. Amaya allows users to specify annotation servers to be checked for annotations and a local directory for local annotations. As new Web pages are browsed, the browser checks the local directory or queries the annotation servers to request any annotations for the current document (HTTP is used to retrieve the annotations from annotation servers). In order to save time and bandwidth, only the annotation metadata is downloaded initially. If the user opens the annotation, the body of the annotation is downloaded. Placement of the annotations in the display is accomplished through the use of XPointer information that indicates where in the document the annotation icon should be placed. An XLink points to the annotation itself. In this way, the annotations can be stored separately from the documents, but the display can be aligned to the correct position. If a document changes in such a way that the anchor for an annotation is removed (i.e. the XPointer is no longer valid), the annotation is considered "an orphan" and can be repositioned [KKP+01, p.627-628]. Other Web annotation systems include similar functionality to Annotea although each system has particular differences.

2.9.3   YAWAS

YAWAS (Yet Another Web Annotation System) [DV00] is a Web annotation tool that uses Level 2 DOM and Dynamic HTML to highlight annotated text on a Web page. The corresponding annotations are displayed in a separate window. Annotations are stored locally in YAWAS and the system includes a search interface to refind annotations. YAWAS works with Microsoft Internet Explorer.

2.9.4   iMarkup

iMarkup is a commercially available system that allows users to make text annotations as well as drawing on and highlighting sections of Web pages [Ima02]. iMarkup offers tools for personal and collaborative annotation use. iMarkup works with Microsoft Internet Explorer.

2.9.5   Microsoft Web Discussions

Microsoft has developed a tool called Web Discussions that integrates with its Internet Explorer Web browser to allow groups to annotate Web pages and conduct asynchronous collaboration. Researchers at Microsoft Research have explored algorithms for positioning annotations [BBG+01] and have studied use of the collaborative aspects of Web Discussions in a study of its use by a Microsoft product development group [CGG00].

2.9.6   IBM WBI

"Intermediaries are software programs or agents that meaningfully transform information as it flows from one computer to another" [MB00, p.96]. For example, a Web proxy can be configured as a type of Web

intermediary that caches Web page accesses for a group of users and attempts to speed up future page requests by servicing them from its local cache instead of re-accessing the pages from a remote server. IBM has developed a powerful set of Web intermediary tools called Web Browser Intelligence ("WBI, pronounced 'WEB-ee'" [BMK97, p.75]) that have been used to create a number of applications that intercept and modify Web pages (and HTTP requests) between Web servers and the user's Web browser [BMK97] [MB00]. WBI has a number of uses, including abilities to "observe user actions, proactively offer assistance, modify Web documents, and perform new functions" [BMK97, p.75]. Using WBI, Web pages can be automatically annotated with additional information. For example, a Web page could be annotated by a "Dictionary plugin" that "turns ordinary words on Web pages into hyperlinks that point to their definitions" [MB00, p.101].

2.9.7   Web Annotation Implementation Issues

As pointed out in [KKP+01], there are two basic approaches to implementing annotations: proxy-based and browser-based. Proxy-based implementations use an HTTP proxy to integrate the annotations with the Web page prior to the receipt of the HTML at the browser. An advantage of proxy-based implementations is that they will work with existing browsers that support the use of proxies. Browser-based implementations handle the integration of the annotations within the browser and have opportunities for greater integration of the annotations into the user interface of the browser.

Chapter 3

Preliminary Survey – Web Tasks and Frequencies

The main study that is the focus of this dissertation required participants to complete a set of tasks to find information on the Web and then, in a separate session about a week later, to refind the same or similar information. One of the factors that I wanted to examine was how the prior frequency with which a person had performed a task or similar tasks affected their searching behaviors. The tasks included in the main study needed to be representative of real-world tasks that people do frequently and infrequently.

In order to guide the development of representative tasks for the main study, I conducted a preliminary study in the form of an on-line survey to find out what tasks people do *frequently* and *infrequently* on the Web. This section describes the development, administration, and results of the survey.

3.1   Survey Methods

The survey was announced on email lists to graduate students in two programs at Virginia Tech: Computer Science and the Human Factors area of Industrial Engineering. Twenty-nine people responded, mostly graduate students from the two departments that received the email. Participation was voluntary and participants were not offered monetary compensation to participate.

The survey was administered using the *survey.vt.edu* Web site. This site is a service that allows an experimenter to design and present a survey as an on-line Web form that respondents can fill-in and submit. The service automatically records and tabulates responses. For this survey, the service was configured so that only people with a Virginia Tech PID (i.e. current faculty, staff, and students) could respond.

The survey asked respondents to list five tasks they do frequently on the Web and five additional tasks that they do infrequently on the Web. "Frequently" was defined in the survey as "daily or weekly." "Infrequently" was defined as "less often than daily or weekly." The survey then asked respondents to recall what steps they take to accomplish six of the tasks they listed (three of the frequent tasks and three of the infrequent tasks). The complete text of the survey is presented in Appendix C. The survey was approved by the Virginia Tech Institutional Review Board with approval #03-371 on August 11, 2003. A copy of the IRB approval memo is presented in Appendix F.

3.2   Survey Analysis

Twenty-nine respondents completed the survey. Respondents listed an average of 4.28 frequent tasks (SD = 0.80) and 3.17 infrequent tasks (SD = 1.42). Cumulatively, the respondents listed a total of 124 frequent tasks and 92 infrequent tasks (a total of 216 task descriptions).

In order to understand the types of tasks the respondents listed, each task description was categorized into one of 15 categories. Rather than using a pre-defined set of categories, I created the category list through a data-driven process that involved reviewing the task descriptions and finding topics that emerged from the task descriptions. Table 1 shows the 15 categories along with examples of actual task descriptions from the survey. One category, "finding information," was further divided into 14 second-level subcategories to indicate the type of information that the respondent was trying to find. The 14 subcategories are shown in Table 2. I used these categories and subcategories to categorize each of the task descriptions.

Another member of my research group was recruited as a "coder" to independently categorize each of the task descriptions using the same set of categories. The initial agreement on top-level categories between the coder's categorizations and mine was 93% (201 out of 216 items)[7]. Of the 15 task descriptions without categorization agreement, 6 were attributed to me mis-coding an item and were fixed. The remaining 9 task descriptions were items that could easily be interpreted as belonging in more than one category. For example, the task description "classifieds", was categorized as "finding information" by either me or the coder and "news/blogs/homepages" by the other. For the 9 task descriptions without agreement, I resolved which of the two categories to use for the final categorization.

After resolving the top-level categories, the initial agreement between me and the coder on the second-level categories for "finding information" was 87% (39 out of 45). Four of the six disagreements were because there were two closely related categories, "computer" and "programming/technical". I decided to combine these categories into "computer/technical". I resolved the two remaining second-level category disagreements.

---

[7] One task description appeared to mention two tasks. The coder submitted two task descriptions for this task and thus submitted 217 categorizations. I reviewed the items and decided that the task description was focused on one activity and removed the additional description from the coder.

Example Task Descriptions from the survey data	Top-Level Category	# of responses Frequent	# of responses Infrequent	# of responses Total
read news blog viewing world news look at personal pages of friends and family	Read news / blogs / homepages	29	5	34
check movie listings get medical information job search	Find information	22	23	45
research papers search for articles	Academic research	18	2	20
checking email Webmail	Email	16	1	17
googling search	Search	9	0	9
order food foreign language translation	Other Web activity	5	6	11
design Web pages post new information to personal Website	Update Website/blog	5	4	9
play games read about hobbies read online comics	Leisure / entertainment	5	3	8
make purchases ebay buying online	Shopping	4	20	24
weather check the weather forecast	Check weather	4	1	5
online banking financial stuff	Banking/accounts/financial	3	9	12
directions getting travel directions	Maps / directions	3	5	8
downloading software look for software	Download software	1	3	4
chatting IM	Instant messaging	0	4	4
format hard disk edit bookmarks	Not applicable – not a Web task	0	6	6
	TOTALS	**124**	**92**	**216**

Table 1. Survey Data Example Task Description, Top-Level Categories, and Frequency Counts

Example Task Descriptions from the survey data	Second-Level Category	# of responses
Solution to computer related problem Read Mac Web sites Software manuals	computer / technical	7
Look up movies Look up stuff to do	entertainment	7
Read instructions for other tasks I'm doing	other	6
Phone numbers Look up contact info	phone / address	4
Check on VT information	Virginia Tech	4
Locate people	people	3
Look up words in a dictionary	dictionary	2
Job search	jobs	2
Search for medical information	medical	2
Reading about products that I intend to buy	products	2
restaurants	restaurant	2
Places to visit	travel	2
Sports update	sports	1
Read stock information	stocks	1

Table 2. Subcategories of Finding Information

The next step in the analysis was to generate frequency counts of the tasks along two dimensions: 1) was the task listed as a frequent or an infrequent task, and 2) the categorization of the task. These frequency counts are shown numerically in Table 1 and graphically in Figure 2.
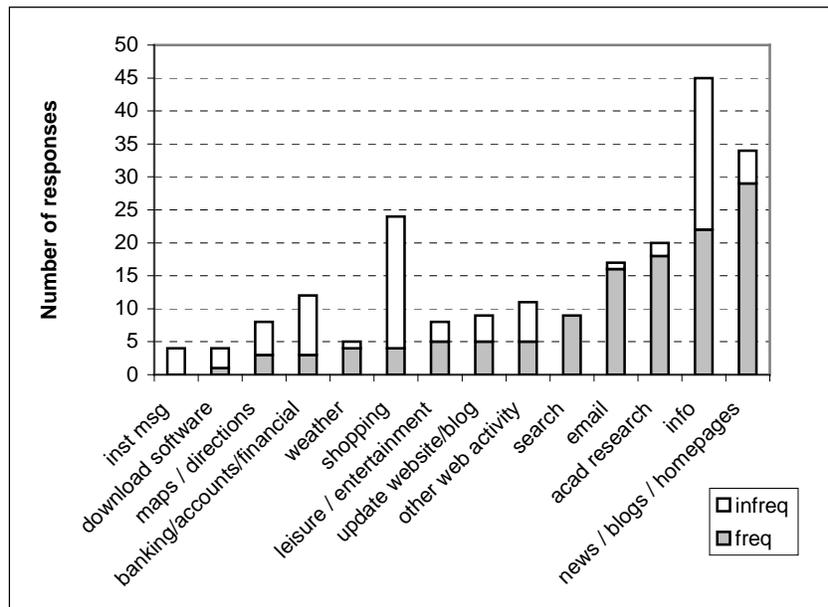
Figure 2. Graph of Frequency Counts for Categorized Survey Task Descriptions

3.3   Survey Results and Discussion

The frequency counts show that some tasks were predominantly listed as frequent tasks, while others were predominantly infrequent tasks. Among the frequent tasks were: reading news/blogs/homepages, finding information (also listed as an infrequent task depending on the type of information), academic research, email, and search. The infrequent tasks included: finding information (depending on the type of information), shopping, and managing on-line accounts such as banking and financial services.

The tasks reported by the respondents to my survey (conducted in 2003) have many overlaps with recent data gathered by the Pew Internet and American Life project about what tasks Internet users have ever performed [Pew05b] and what tasks Internet users do on a typical day [Pew05a].

The survey method used in my study did not prompt users with task descriptions. Instead, it asked users to recall both frequent and infrequent tasks they performed on the Web. There are both benefits and drawbacks to this approach. Respondents were not limited in their choices of tasks, but were limited to providing up to five tasks for each area of interest – frequent and infrequent tasks. The free-response nature of the questions may elicit a broader range of tasks than rating to a fixed set of tasks. However, free response relies on the respondent's memory and they may forget important tasks. In addition, the tasks listed may be biased toward tasks that they easily recall. Thus, the infrequent tasks may be ones that are infrequent, but memorable. Infrequent tasks that were not recalled would not be listed, even if they are important. The fact that most of the tasks gathered in the survey overlap with tasks in the Pew Internet and American Life studies [Pew05a] [Pew05b] provides good confirmation that these tasks are representative and are tasks that are of interest for finding and refinding research.

The tasks and data gathered from the survey were used to help inform the design of a set of 18 tasks to be used in the main study of finding and refinding presented in this dissertation. Table 3 shows how the tasks included in the study relate to the tasks identified by the survey. Not all the tasks identified in the survey could be included in the main study, but there is good coverage, especially in tasks to find information. Descriptions of the 18 tasks are presented in Appendix E.

Category / Subcategory	Task Created for use in main study	Task # in main study
Read news / blogs / homepages	News – read a news story about world news	2
Find information		
Dictionary	Dictionary – look up meaning of a word	1
Phone / address	Phone – local art supply store	3
Travel	Travel – flight reservation	5
Other	Other – Rubik's cube solution	6
Travel, Entertainment	Travel – tourist attractions in San Francisco	7
Phone / address	Phone – local grocery store	8
Sports	Sports – NFL football score	10
Entertainment	Entertainment – movie showtimes	11
Stocks	Stocks – current price of Microsoft stock	12
Other	Other – instruction manual for VCR	15
People, Academic	People – Web site of Dr. Stasko at Georgia Tech	13
VT	VT – last day of classes	17
computer / technical		n/i
jobs		n/i
medical		n/i
products		n/i
restaurant		n/i
Academic research		n/i
Email		n/i
Search		n/i
Other Web activity		n/i
Update Website/blog		n/i
Leisure / entertainment		n/i
Shopping	Shopping – women's sweatshirts for a gift	4
	Webpage – go to the Wal-Mart Web site	9
Check weather	Weather – hi and low in Madrid, Spain	14
	Weather – hi and low in Blacksburg, VA	16
Banking/accounts/financial		n/i
Maps / directions	Map – directions / distance between two locations	18
Download software		n/i
Instant messaging		n/i
Not applicable – not a Web task		n/i

* n/i indicates that a survey task was not included in the main study

Table 3. Coverage of the survey tasks in the main study

Chapter 4

Method

This section describes the main study of this dissertation – a controlled, laboratory study of how people find and refind information. This study had two experimental sessions, scheduled about a week apart. In the first session, participants were given a set of 18 tasks that involved finding information on the Web. In the second session, participants were given a set of 18 counterpart tasks, most of which involved refinding information that had been found in the first session.

4.1 Participants

Participants in this study were all undergraduate students at Virginia Tech. Recruitment was conducted widely across campus to insure a variety of backgrounds and majors. A total of 21 people participated, but only 17 completed both sessions with usable data on both days. The results presented here are based on these 17 participants. Participants were offered a payment of twenty dollars for their participation. This study was approved by the Virginia Tech Institutional Review Board with approval #03-539 on November 12, 2003. A copy of the IRB approval memo is presented in Appendix F.

4.2 Materials and Equipment

This section describes the materials and equipment used to conduct the study.

4.2.1 Location, Equipment and Data Collection

The study was conducted in the Virginia Tech Department of Computer Science Usability Evaluation Lab in McBryde 102. Participants completed the main part of the study in a participant room (McBryde 102A) that was separated from the experimenter control room (McBryde 102B) by a one-way mirror. Participants used a personal computer running Windows 2000 and Microsoft Internet Explorer to complete the tasks. The experimenter was able to observe participants as they worked on the tasks through the one-way mirror. The lights in the control room were kept dim so that the participant could not see into the control room.

Remote desktop monitoring and control software (RealVNC, www.realvnc.com) was installed on both the computer in the participant room and on the computer in the control room. This software sent a duplicate screen image of the participant computer to the control room computer so that the experimenter could observe in real-time the actions of the participant. The VNC software also allowed the experimenter to remotely control the computer in the participant room. This feature was very useful in situations where a participant accidentally clicked the "next" button on the data collection program – the experimenter could take control of the participant computer while still sitting in the control room and get the program back to the correct task.

The rooms were also equipped with recording devices/software and an intercom Screen and audio recording software (Camtasia, www.techsmith.com) was installed on the participant computer. This software created a movie file (.avi) that was a recording of the screen as the participant worked on the tasks. A microphone was placed on the participant room desk and its output was connected to the screen recording system so that any verbal comments that the participant made would be captured along with the screen recording. Output from this microphone was also connected to speakers in the control room so that

the experimenter could always hear the participant with no need for the participant to press any buttons. A press-to-talk intercom system was configured with units in both rooms. The unit in the participant room was located in the back left corner of the desk (out of the way) and used mainly as a speaker so that the participant could hear communications from the experimenter. The experimenter used the push-to-talk button on the unit in the control room to enable communication to the participant. Using this configuration, the experimenter could always hear the participant (through the sound recording microphone), but the participant could only hear the experimenter when the intercom was used.

In addition to the screen and audio recording software, a commercial logging program (Spector, www.spectorsoft.com) was installed on the computer in the participant room. This program logged URLs that were requested and keystrokes that were typed.

A custom data collection program was written in Visual Basic to run as a stand-alone program that would present the set of 18 tasks to the participants during each experimental session and collect data from the participant about each task. Screen shots from the program are shown in Figure 3 and Figure 4 (a more extensive set of screen shots of the VB program is presented in Appendix D).

Using the URL log, the data from the VB program, and the screen recordings, the following measures were recorded and calculated for each task: task completion time, number of URLs loaded, the number of uses of Web search engine results, the ratio of search engine uses to the total number of pages loaded, and information gathered from the participant about their experience with each task (more details on this in section 4.4.4).

4.2.2 Demographic Questionnaire

During the first session, each participant completed a demographic survey. This survey was administered using an on-line survey tool called *survey.vt.edu* that displays survey questions as a Web page form in a Web browser. Most participants completed the survey using the computer in the participant room. However, due to time constraints, one or two participants completed the survey using a different computer in the McByrde 102 main lab area while they waited for another participant to finish in the participant room.

4.2.3 GEFT and Cube comparisons

On each day, prior to starting the 18 tasks, participants were given a written test to measure cognitive aspects of interest. During the first session, participants were given the Group Embedded Figures Test (GEFT) [WOR+71] to measure field dependence/independence. During the second session, participants were given the Cube Comparisons test (S2) from the Kit of Factor Referenced Cognitive Tests [EFH+76] to measure spatial orientation. The results of these tests are not included in the analysis presented in this dissertation.

4.3 Procedure

Participants completed two experimental sessions, scheduled about a week apart. Each session is described in more detail below.

4.3.1 Session 1

Participants were greeted by the experimenter and given an informed consent form to read and sign. The experimenter then explained and administered the Group Embedded Figures Test. Following this, the participant was seated at a computer and asked to complete the on-line demographic questionnaire. After completing the questionnaire, the experimenter started the Visual Basic program that presented the

participant with the 18 tasks to find information on the Web. For each participant, the tasks were presented in a random order.

Participants were given approximately 45 minutes to work on the 18 tasks at their own pace. However, if a participant got stuck on a particular task, the experimenter might suggest to them that they could give up on the task and move on to the next one. When the session was completed, the experimenter thanked the participant for their time and reminded them of the date and time of their second session.

4.3.2   Session 2

The experimenter greeted the participant and then explained and administered the Cube Comparisons (S2) test from the Kit of Factor Referenced Cognitive Tests [EFH+76]. Then, the participant was seated at the same computer that they were at in the first session and were given another set of 18 tasks that involved locating information on the Internet using a Web browser. Some of the tasks in the second session were exactly the same as the session one tasks, while others differed slightly. The experimenter again started the Visual Basic program to present the 18 tasks to the participant. For each participant, the tasks that they had completed in session 1 were presented in a new random order and were then followed by the remaining tasks (the ones they had not completed in session 1), presented in the same random order that was generated for session 1.

Participants were given approximately 45 minutes to work on the 18 tasks at their own pace. However, if a participant got stuck on a particular task, the experimenter might suggest to them that they could give up on the task and move on to the next one. When the session was completed, the experimenter thanked the participant, paid them for their time, and had them sign a paper acknowledging receipt of their payment[8].

---

[8] One participant was not offered payment until after they had finished and left the experiment. They were contacted by email and offered payment, but never collected it.

4.3.3   Tasks

In each session, a series of 18 tasks were presented to the participants through a custom-designed Visual Basic (VB) program. For each of the 18 tasks, the VB program presented the following sequence of three screens: before-task questions, task information collection, and after-task questions. Each of these is described in more detail below. Note that the task description shown in these images is a sample task that was not used in the actual study.

- *Before task questions*. The VB program displayed a brief description of the current task along with a set of questions to measure aspects of the participant's prior experience with this or similar tasks. Figure 3 shows an example of a before-task questions screen.

Figure 3. Example of a Before-Task Questions Screen

- *Task information collection*. Each task consisted of a task description and a set of target information to be found on the Web. For example, one task description was to "Find tomorrow's hi and lo temperature for Blacksburg" and the target information was the high and low temperature. The VB program displayed a screen with the task description and a set of text entry fields (or radio boxes depending on the task) for the participant to enter the

information requested by the task. A complete list of the task descriptions and the information requested for each task is presented in Appendix E.

▪ *After task questions*. After the participant completed the task and clicked on "next", the VB program displayed a set of after-task questions to inquire about aspects of the participant's experience with the task. Figure 4 shows an example of an after-task questions screen.

Figure 4. Example of an After-Task Questions Screen

For each task, the participant was to answer the before task questions, then search for the information requested by the task, and then answer the after task questions.

4.4   Experimental Design

This section describes the experimental design for the main study of finding and refinding. Descriptions of the overall design and each of the dependent and independent variables are given. At the end of the section, specific experimental hypotheses are presented for each of the research questions introduced in section 1.3.

The main study design is a 2x18 within-subjects design with factors Session (1, 2) and Task (1..18), and Participant as a repeated measure. All participants were given the same set of tasks. Tasks were either exactly the same or similar from session 1 to session 2, with the exception of two tasks that were changed in session 2 (more information is given on the relationship of tasks between session 1 and session 2 in section 4.4.2.5). The same measurements about the tasks were made in session 1 and session 2. A number of dependent measures were collected, some of which were used to group tasks in additional analyses (see section 4.4.4).

4.4.1   Complexities in the Study

Studying differences between finding and refinding can be difficult to study, especially in a laboratory environment. For example, a finding task for one participant may be a refinding task for another. A primary issue is how to define what constitutes the difference between a finding task and a refinding task. In this dissertation, I operationalize the finding versus refinding task distinction in two ways: 1) based on Session, and 2) based on Task Frequency. Each of these is described below.

In this study, one operationalization of the finding/refinding distinction is based on Session and relies on the study design that 16 of the 18 tasks given in the second session were similar to or subsets of tasks given in the first session. For all the participants who completed a task during the first session, the counterpart task in the second session would be a refinding task. Thus, Session acts as a gross measure of whether a task is a finding task (session 1) or a refinding task (session 2). This measure has the advantage that it is a factor that was controlled in the study design. The "Controlled Factors Analysis" in section 5.2 considers finding and refinding using this measure.

However, using session to distinguish finding and refinding tasks does not fully account for participants' prior experience doing tasks on the Web. For example, many people check the local weather on the Web every day, so this task would not be a finding task for them, even in session 1. In the before-task questions, participants were asked to rate their prior frequency with each task. This self-reported measure of prior frequency is the basis of the second method used in this dissertation to distinguish finding and refinding tasks. Tasks with high prior frequency can be considered refinding tasks and tasks with very low or no prior frequency can be considered finding tasks. This illustrates the concept that finding and refinding lie on a continuum based on frequency. The "Prior Frequency Analysis" presented in section 5.3 considers finding and refinding as distinguished by prior frequency. Tasks are divided into groups based on high, medium, and low levels of prior frequency (frequency measured before doing the task in session 1) and differences are examined between these three groupings.

The classification using prior frequency is most relevant for the tasks completed in session 1 because it is based on the participant's *prior* frequency of doing the task, before participating in the study. In session 2, the participant has already done the tasks once and has the benefit of that experience, so their prior frequency may have less of an impact. For this reason, the "Prior Frequency Analysis" in section 5.3 contains two analyses – one analysis that considers data from both session 1 and session 2 based on prior frequency, and a second analysis that considers data only from session 1 based on prior frequency (see section 5.3.4).

4.4.2   Independent variables

There were two main independent variables in the study: Session and Task. In addition, analyses were conduced by grouping tasks based on *task frequency*, *task familiarity*, and *task similarity*. The independent variables and the groupings are described in more detail below.

4.4.2.1   *Session*

Session is the experimental session in which the data were collected – either the first session (session 1), or the second session (session 2) which was scheduled about a week after the first session. A total of 17 participants completed tasks with usable data in both sessions.

4.4.2.2   *Task*

Participants were presented with 18 tasks in both sessions. Task development was informed by the survey of frequent and infrequent tasks that is reported on in Chapter 3. All but two of the tasks presented in the second session concerned refinding information related to information requested in the tasks for the first session. Two tasks were changed from the first session to the second session – these changed tasks are not considered in the analyses presented here. A complete listing of the 18 tasks used in session 1 and session 2 is presented in Figure 5 and in more detail in Appendix E.

4.4.2.3   *Task frequency*

Task frequency data was gathered by a before-task question for each task that asked participants, "How often do you do this type of task using the Internet in your day-to-day life?" Responses available were: more than once per day, once per day, several times per week, once per week, several times per month, once per month, less often than once per month, and never.

When this variable was used to group tasks, the groupings were:

- High = once per week or more
- Medium = less than high but at least once per month
- Low = less often than once per month, or never

4.4.2.4   *Task familiarity*

Task familiarity data was gathered by a before-task question for each task that asked participants, "How well do you know how to get to the information?" The response scale was a 7-point, equally spaced, Likert-type scale, with anchors, "I Don't know where it is" = 1, and "I Know exactly where it is" = 7.

When this variable was used to group tasks, the groupings were:

- High = 7 or 6
- Medium = 5, 4, 3
- Low = 2 or 1

Figure 5. Tasks for Session 1 and Session 2

Task Num		Session 1	Similarity	Session 2
1	Dictionary	Look up the meaning of the word, salubrious	Path	Look up the meaning of the word, frugivorous
2	ReadNews	Find and READ news story about world news	Moved	Find the news story that you read last week about world news.
3	Phone	Find the phone number for the Mish Mish art supply store in downtown Blacksburg	Exact	Find the phone number for the Mish Mish art supply store in downtown Blacksburg
4	Shopping	Look for prices on Women's Sweatshirts. Find two possibilities that you might buy for a gift to a friend or relative. Please look at at least three different sweatshirts from three different places to buy.	Subset	Find one of the Women's Sweatshirts that you looked at during the session last week.
5	Travel	Find a flight from Roanoke, VA on Dec. 1 to Chicago, IL and return on Dec. 4	Path	Find a flight from Roanoke, VA on Dec. 14 to Chicago, IL and return on Dec. 19
6	Research	Find instructions about how to solve the Rubik's cube.	Exact	Find instructions about how to solve the Rubik's cube.
7	Travel	Imagine that you are planning a trip to San Francisco, California. Find two tourist attractions or events that you might be interested in. Please look at at least four different attractions or events.	Subset	Find one of the tourist attractions or events for San Francisco, California that you looked at during the session last week.
8	Phone	Find the phone number for the Kroger grocery store in Blacksburg near the Math Emporium	Path	Find the phone number for the Kroger grocery store in Blacksburg on South Main Street near El Guadalupe
9	Homepage	Go to the Wal-Mart Web site	Different	Go to the Web site for Dr. David Patterson at the University of California at Berkeley
10	Sports	Find the score from the most recent Atlanta Falcon's NFL football game.	Exact	Find the score from the most recent Atlanta Falcon's NFL football game.
11	Movies	Find a movie that you are interested in seeing. Find a list of the starting times for the movie for today.	Exact	Find a movie that you are interested in seeing. Find a list of the starting times for the movie for today.
12	Stocks	Find the current price for Microsoft stock.	Exact	Find the current price for Microsoft stock.
13	Homepage	Go to the Web site of Dr. Stasko at Georgia Tech	Exact	Go to the Web site of Dr. Stasko at Georgia Tech
14	Weather	Find tomorrow's hi and low temperature for Madrid, Spain	Different	Go to the Web site for The Record Exchange music store on Main Street in Blacksburg
15	Research	Find an instruction manual for a Panasonic AG-1340 Video Cassette Recorder (VCR)	Exact	Find an instruction manual for a Panasonic AG-1340 Video Cassette Recorder (VCR)
16	Weather	Find tomorrow's hi and lo temperature for Blacksburg	Exact	Find tomorrow's hi and lo temperature for Blacksburg
17	Academic	Find out what the last day of classes is for the Spring 2004 semester	Path	Find out when Spring Break is for the Spring 2004 semester
18	Map	Find out how to get from the Blacksburg Municipal Golf Course to the Macados on University Blvd.	Exact	Find out how to get from the Blacksburg Municipal Golf Course to the Macados on University Blvd.

4.4.2.5   Task similarity

Task similarity is the similarity of a task from the first session to the second session. Four types of task similarity are considered in this study: exact, path, subset, and moved. Each of these is described below and an example is given. There is more discussion of task similarity and related concepts in section 7.2.1.2 of this dissertation.

> Exact (E) – exactly the same task on both days
>> e.g.:   Session 1:   Find the phone number for the Mish Mish art supply store.
>> Session 2:   Find the phone number for the Mish Mish art supply store.

> Path (P) – the session 2 task information was on the same path as the session 1 information
>> e.g.:   Session 1:   Find the phone number for the Kroger grocery on *South Main St.*
>> Session 2:   Find the phone number for the Kroger grocery on *University Blvd.*

> Subset (S) – the day two task information was a subset of the information found on day one
>> e.g.:   Session 1:   Find two sweatshirts that you might buy as a gift for a friend.
>> Session 2:   Find *one* of the sweatshirts that you found during the session last week.

> Moved (M) – the day two task was the same as day one, but the information moved
>> e.g.:   Session 1:   Find and READ a news story about world news.
>> Session 2:   Find the news story *that you read last week* about world news.

4.4.3   Dependent variables – calculated/measured

Data for a number of dependent variables were recorded and calculated from the log files and screen recordings for each participant performing each task in each session. These are described in the following sections.

4.4.3.1   Task completion time

Task completion time was measured as the time from when the participant initiated the task by clicking "next" in the data collection program at the end of the before-task questions until they clicked "next" to go to the after-task questions. Task start and end times for each task were coded manually by one of two coders based on watching the screen recording movies. I was one of the coders and the other coder was a trained HCI expert. Adjustments to the task completion times were made in cases where there was an interruption or delay. For example, in several cases, a participant accidentally clicked the "Next" button on the Visual Basic data collection program and the experimenter had to reset the program to the correct place. In these cases, the time taken by the experimenter to reset the program was subtracted from the task completion time for that task.

4.4.3.2   Number of URLs

A data-logging program (Spector) generated a log file of URLs that were loaded/requested as the participants browsed Web pages. These log files were used to generate a count of URLs loaded for each task for each participant.

The program that logged the URLs added a new entry every time a page was requested. This meant that when a participant clicked the "Back" button on the Web browser, the page they went back to would be recorded in the log a second time. For example, if a participant started at www.vt.edu, clicked a link to go to www.vt.edu/students and then clicked the "Back" button on their browser, then the URL logging program would record the following sequence:

www.vt.edu
www.vt.edu/students
www.vt.edu

It was decided that counting each re-accesses to the search engine results pages was appropriate for this study because the participants were potentially referring back to the search results as part of their finding or refinding process.

Two issues should be noted about the data logged by the logging program:

- The logging program recorded all URL requests, including pop-up and pop-under advertisements. For the analyses presented here, no attempt was made to filter these from the log files.

- The logging program would occasionally miss recording certain URLs. For example, the log files are often missing entries for the main page of yellowpages.com even though participants visited this URL. These errors seem to be inherent in the logging program and are believed to have affected only a few URLs and are also believed to have had little effect on the results.

4.4.3.3   *Number of Searches*

The number of uses of Web search engines was counted. This was done using a custom written Perl script that scanned the URLs in the log file for each task. The script counted a URL as search engine results page if it contained any of the following strings:

- `search.yahoo.com`
- `altavista.com/Web/results`
- `google.com.results`
- `search.msn.com`

At the time of the study, these strings identified result pages returned from major Web search engines that participants were observed using in the study.

4.4.3.4   *Search Ratio*

Since the tasks were of different lengths, I developed a search ratio measure to represent the use of Web search engines results normalized based on the total number of URLs. In this study, the measure, *search ratio,* is defined (and computed) to be the total number of searches divided by the total number of URLs loaded for the task.

4.4.4   Dependent variables – response

Data for a number of dependent variables was gathered from the VB program for each participant performing each task in each session. These are described in the following sections.

4.4.4.1   *Task Frequency*

Task frequency data was gathered by a before-task question for each task that asked participants, "How often do you do this type of task using the Internet in your day-to-day life?" Responses available were: more than once per day, once per day, several times per week, once per week, several times per month, once per month, less often than once per month, and never.

Important Note: In the analyses presented in Chapter 5, when used as a dependent variable, task frequency is often treated as interval data on a 1 to 8 scale although it was measured on an ordinal scale. This limits the comparisons and conclusions that can be drawn about the data.

4.4.4.2   Task Familiarity

Task familiarity data was gathered by a before-task question for each task that asked participants, "How well do you know how to get to the information?" The response scale was a 7-point, equally spaced, Likert-type scale, with anchors, "I Don't know where it is" = 1, and "I Know exactly where it is" = 7.

4.4.4.3   Perceived task difficulty – before and after task

Perceived before task difficulty data was gathered by before- and after-task questions for each task that asked participants, "How difficult do you think this task will be?" (before task), and "How difficult was this task?" (after task). For both questions, the response scale was a 7-point, equally spaced, Likert-type scale, with anchors: very difficult (1), difficult (2), somewhat difficult (3), average (4), somewhat easy (5), easy (6), and very easy (7). Note that the numeric assignment is from difficult (1) to easy (7), so that higher values indicate easier ratings.

4.5   Hypotheses

In this section, specific hypotheses for each research question are described.

4.5.1   RQ1 – Differences between finding and refinding

RQ1 – Do finding and refinding differ? Specifically, are there measurable differences between finding and refinding in terms of user knowledge/familiarity, perceived difficulty, search engine use, and task length?

This research question is examined using two ways of defining finding and refinding tasks. In the first method, session 1 tasks are considered finding tasks and session 2 tasks are considered refinding tasks. Hypothesis 1 (see below) is focused on the first method. In the second method, prior frequency with a task is considered as a measure of where a task lies on a continuum from finding to refinding. Tasks are divided into groups of high, medium, and low prior frequency based on participants' responses in session 1 to the before-task question about prior frequency with the task. Hypothesis 2 is focused on the second method.

> *Hypothesis 1 – Finding and Refinding have differences when defined by Session*
> *There will be a significant effect of Session on task completion time, number of URLs, number of searches, search ratio, familiarity, before-task difficulty, and after-task difficulty, but not on frequency. Specifically, task completion time, number of URLs, search ratio, before-task difficulty, and after-task difficulty will all be lower/easier for Session 2 than for Session 1. Familiarity will be higher for Session 2 than for Session 1.*

> Analysis: Analysis 1 – Controlled Factors Analysis (Section 5.2, page 54)
> Results: RQ1 (Section 6.1, page 113)

> *Hypothesis 2 – Finding and Refinding have differences when defined by Task Frequency*
> *There will be a significant effect of task frequency on task completion time, number of URLs, number of searches, search ratio, familiarity, before-task difficulty, and after-task difficulty for the Session 1 and Session 2 tasks. Specifically, task completion time, number of URLs, number of searches, search ratio, before-task difficulty, and after-task difficulty will all be lower/easier for tasks with a high level of frequency than for tasks with a low frequency. Familiarity will be higher for high frequency tasks than for low frequency tasks.*

> Analysis: Analysis 2 – Prior Frequency Analysis (Section 5.3, page 71)
> Results: RQ1 (Section 6.1, page 113)

4.5.2   RQ2 – Task Differences

RQ2 – Do task and familiarity with a task affect finding and refinding? What types of tasks are easy and which are difficult? Specifically, how do task and task familiarity (knowledge of how to perform the task) affect user knowledge/familiarity, perceived difficulty, search engine use, and task length?

This research question examines the effect of task and task familiarity on finding and refinding. Hypotheses 3 and 4 address this research question.

> *Hypothesis 3 – Task has a large effect on finding and refinding*
> *There will be a significant effect of Task on the dependent measures task completion time, number of URLs, number of searches, search ratio, familiarity, frequency, before-task difficulty, and after-task difficulty for the Session 1 and Session 2 tasks. Post-hoc tests will indicate the effects on a per task basis.*

>> Analysis: Analysis 1 – Controlled Factors Analysis (Section 5.2, page 54)
>>   Analysis 6 – Subset and Moved Tasks Analysis (Section 5.6, page 100)
>> Results: RQ2 (Section 6.2, page 117)

> *Hypothesis 4 – Task familiarity affects finding and refinding*
> *There will be a significant effect of prior task familiarity on the dependent measures task completion time, number of URLs, number of searches, search ratio, frequency, before-task difficulty, and after-task difficulty for the Session 1 and Session 2 tasks. Specifically, task completion time, number of URLs, number of searches, search ratio, before-task difficulty, and after-task difficulty will all be lower/easier for tasks with a high level of prior familiarity than tasks with low familiarity. Frequency will be higher for high prior familiarity tasks than for low prior familiarity tasks.*

>> Analysis: Analysis 3 – Prior Familiarity Analysis (Section 5.4, page 83)
>> Results: RQ2 (Section 6.2, page 117)

4.5.3   RQ3 – Finding to Refinding Relationship

RQ3 – How does the relationship of the refinding task to the finding task affect the refinding process? Specifically, does it affect user knowledge/familiarity, perceived difficulty, search engine use, and task length?

This research question explores the effect of the similarity of the task from finding to refinding. In this study, four relationships between finding and refinding tasks are examined: 1) *exact* – tasks are exactly the same in both sessions, 2) *path* – the refinding task involves looking for information on the same path as the information that was first found, 3) *subset* – the refinding task involves looking for a subset of information found, and 4) *moved* – the refinding task involves looking for information that has moved since it was found. Hypothesis 5 addresses this research question.

Hypothesis 5 – Task similarity affects refinding
There will be a significant effect of task similarity on the dependent measures task completion
time, number of URLs, number of searches, search ratio, frequency, before-task difficulty, and
after-task difficulty on the Session 2 (refinding) tasks. Specifically, task completion time, number
of URLs, number of searches, search ratio, before-task difficulty, and after-task difficulty will all
be higher for moved and subset tasks than for exact or path tasks. No difference will be observed
for any of the dependent measures between the exact and path tasks. In other words: Refinding
will be easier if the relationship of the finding and refinding tasks are "exactly the same" or "on
the same path". Refinding will be more difficult if the relationship is "subset" or "moved".

Analysis: Analysis 5 – Task Similarity Analysis (Section 5.5, page 93)
Results: RQ3 (Section 6.3, page 120)

4.5.4   RQ4 – Task Details and Strategies

RQ4 – What strategies are used in finding and refinding? Are there observable patterns of strategy use across types of tasks or from finding to refinding for a single task?

This research question is addressed through a task-by-task analysis of each task to explore if there are common patterns and strategies used.

Analysis: Analysis 8 – Task-by-Task Analysis (Section 5.7, page 101)
Results: RQ4 (Section 6.4, page 122)

This page intentionally left blank.

Chapter 5

Results

This chapter presents the results of analyses performed on the data collected in the main study described in Chapter 4.

Of the 18 tasks in the study in the main analyses, only 12 are typically included. Some of the tasks could not be fairly compared across sessions for many of the dependent variables.

- Two tasks were completely different from session 1 to session 2 and are not included in the statistical analyses.
- The two subset tasks required participants to find two items in session 1, but to refind only one item in session 2. Thus, comparing task completion time for the subset tasks across the two sessions is not useful in a statistical sense.
- The one moved task was only successfully completed by two participants in session 2 and is not included in any of the statistical analyses. It is considered separately in the analysis of the subset and moved tasks in section 5.6.
- One of the exact tasks, task 15 (locating a specific VCR manual) was verified as successfully completed in session 1 by only five participants. Due to the low successful completion rate in session 1 it is not included in the statistical analyses.

The remaining 12 tasks are all the exact and path tasks except for task 15 as noted above. This set is referred to in this dissertation as EP-15 (exact and path tasks minus task 15) and is the primary set of tasks that are used in most of the statistical analyses that involve comparisons across sessions. Of the EP-15 set, eight are exact tasks and four are path tasks. The EP-15 tasks all asked participants to find and refind the same or similar information in both sessions. EPS-15 is a related set that includes EP-15 and the two subset tasks. This group is used in the analysis of task similarity. In addition to the EP-15 label, often the labels S1, S2, and S12 will be used to indicate that an analysis includes only session 1 data (S1), only session 2 data (S2), or both session 1 and session 2 data (S12).

5.1 Data Characteristics

There are several properties and characteristics of the data that are important. Four specific characteristics are described in this section: data exclusion, missing cells, data logging issues, distribution of the data, and analysis methods.

5.1.1 Data Exclusion and Inclusion Criteria

A total of 21 people participated in the study, but only 17 completed both sessions with usable data on both days. The partial data from the other 4 participants was excluded from the analysis.

In all the statistical analyses reported on in this chapter, the data were filtered to only include tasks that were completed and that had usable data. Three main criteria had to be met in order for a task to be included in the statistical analyses: 1) it had to have been *attempted* by the participant, 2) it had to have been *completed* as evidenced by the participant having responded to the Visual Basic (VB) data collection program questions about the task, 3) the participant had to have *responded to the before-task questions*

presented by the VB data collection program before starting the task, and 4) a *manual review* of the screen recording movie had to have not revealed any problems with the task. In addition, in order for a session 2 task to be included in the analysis, all four of the criteria for the session 1 counterpart task had to have been met. Based on these criteria, out of 306 possible total tasks in session 1 (18 tasks x 17 participants = 306 total tasks), 284 were attempted and 264 were included in the analyses. In session 2, out of the 306 possible tasks, 294 were attempted and 229 were included in the analyses.

An overriding principle in deciding whether or not to include a participant's task in the statistical analysis was not to impose a judgment about whether they "got the task right", but rather to assess the question, "did the participant think that they had completed the task?" The following sections describe in detail the three criteria that had to be met to include a task in the analyses.

5.1.1.1   Attempted

Tasks were deemed to have been attempted if the participant reached the after-task questions for that task in the VB data collection program. In order to reach the after-task questions, the participant would have had to have clicked "Next" twice – once to get past the before-task questions and once to get past the questions that requested the task information.

5.1.1.2   Completed

A task was decided to have been completed if the participant provided responses to the questions presented by the VB data collection program about the task. In keeping with the principle of counting a task if the participant thought they had completed it, in many cases, specific responses to the questions were not manually checked for accuracy, but rather, the input fields were checked using a Perl script to see if the participant had provided any response. The presence of an answer was typically taken as enough evidence of completion to include the task in the analyses. One reason behind this methodology was that if a participant thought they had completed the task, then they had – people often find "incorrect" information when searching. This study was primarily focused on studying the processes of finding and refinding. For the subset tasks, it was important to establish the accuracy of performance on the refinding task in order to decide whether to include the task. For the subset tasks, accuracy was an important aspect of making comparisons to other tasks and making assessments of the finding to refinding process. The subset tasks were checked to see if the participant actually re-found one of the items that was found during the first session – the task would not be a refinding task otherwise.

The VB data collection program asked participants two main types of questions regarding the tasks: 1) yes or no questions, and 2) fill-in-the-blank text box(es) to request specific information. Details about each type of task response are given below.

Yes or No – several tasks asked participants to respond simply by indicating "yes" or "no" to whether the requested information had been found. For example, the task to refind the news article asked participants "Did you find it?" and presented radio buttons for a yes or no response. These tasks were included in the statistical analyses if the participant responded "yes" and were not included if the participant responded "no".

Input text box(es) – these tasks requested that participant fill-in text boxes with information found about the task. For example, the task to find the local weather requested that participants type in the high and low temperature for the next day based on the forecast they found. These tasks were included in the statistical analyses if all the required text boxes had some text present. Otherwise, the task was not included.

5.1.1.3   Before-task questions

The participant had to have completed the before-task questions and to have done so before actually starting the task. In a handful of cases, a task had to be excluded because it was noticed in the manual review of the screen-recording movie that the participant started the task before they completed answering the before-task questions. There was nothing in the data collection program to prevent this – it operated

completely independently from the Web browser. The task description had to be displayed along with the before-task questions so that the participant could provide accurate answers about the task. As such, occasionally participants would read the task description at the top of the before-task questions screen and start doing the task before answering the questions. In these cases, the task had to be excluded because the answers provided for the before-task questions would not be valid.

5.1.1.4   Manual Review

While coding the start and end times for the tasks, the coders noted any out of the ordinary events that would require the data for a task to be excluded. A handful of tasks were excluded based on these notes. Many of the tasks excluded because of the manual review were instances where the participant started attempting the task before answering the before-task questions.

5.1.2   Missing cells

Because of the data inclusion and exclusion criteria and because some participants did not attempt all the tasks due to time constraints (participants were given about 45 minutes to work on the 18 tasks at their own pace), the data set has missing cells. This means that in the statistical analysis when the variables Participant and Task are crossed (i.e. Participant x Task), some cells are missing. Special analysis techniques must be used when a data set has missing cells (the analysis techniques are discussed in section 5.1.4).

In session 1, the tasks were presented in a random order for each participant. In session 2, tasks were presented as follows: first, the set of tasks that the participant had completed during session 1 was presented using a new random order; then the set of tasks that the participant had not completed in session 1 were presented using the same random order that was generated for session 1.

Because the tasks were presented to participants in random order, the tasks that were not *attempted* by the participants due to time constraints should be distributed at random. However, the tasks that participants attempted but did not *complete* successfully are probably not completely random. For example, in session 1, only 7 participants completed task 15 (VCR manual). In session 2, only 2 participants completed task 2 (news article). Because of the low completion rate of these two tasks, they have been excluded from the main analyses.

The distributions of the missing cells in each session due to participants not attempting a task are shown in Table 4 (session 1) and Table 5 (session 2). The distributions of the cells that are missing due to the tasks not meeting the criteria specified in section 5.1.1 are shown in Table 6 (session 1) and Table 7 (session 2). In addition to the missing cells as indicated in these tables, there are occasional additional missing cells for particular dependent variables – in some cases a participant completed a task, but did not provide an answer to a specific before or after task question.

Table 4. Distribution of Missing Cells in the Participant x Task Data for Session 1

		Task																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
P a r t i c i p a n t	12	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	0
	13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	14	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0
	15	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0	0
	16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	17	1	0	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0
	18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	20	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	21	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0
	22	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	1	0	1
	23	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1
	25	1	0	1	1	1	1	0	1	1	1	0	1	0	1	0	1	1	0
	26	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1
	29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	31	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
	32	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1

Table 5. Distribution of Missing Cells in the Participant x Task Data for Session 2

		Task																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
P a r t i c i p a n t	12	1	0	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	0
	13	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	14	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0
	15	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0	0
	16	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0
	17	1	0	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0
	18	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	20	1	0	0	1	0	1	1	0	1	1	1	1	1	1	0	1	1	0
	21	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0
	22	1	0	1	0	0	0	1	1	0	0	1	0	0	1	0	1	0	1
	23	1	0	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1
	25	1	0	1	1	1	1	0	1	1	1	0	1	0	1	0	1	1	0
	26	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1
	29	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1
	30	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1
	31	1	0	1	0	1	0	1	1	0	1	1	1	0	1	1	1	1	1
	32	1	0	0	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1

Table 6. Distribution of missing cells due to tasks not attempted for session 1

		Task																	
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
P a r t i c i p a n t	**12**	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
	15	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0
	16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	17	1	0	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0
	18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
	20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	22	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	1	0	1
	23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	25	1	0	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	0
	26	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1
	29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	31	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 7. Distribution of missing cells due to tasks not attempted for session 2

		Task																	
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
P a r t i c i p a n t	**12**	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
	16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	22	1	1	1	0	0	0	1	1	1	0	1	1	0	1	0	1	0	1
	23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	25	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0
	26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	31	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

5.1.3   Data Logging Omissions

In reviewing the data logs collected by the data logging software, some URLs that users visited were not logged correctly. Currently, only a few specific URLs have been found that were affected by this problem. It is believed that these URL log omissions were small in number and that they do not have a large effect on the results. It should also be noted that the URL logs include URLs of any pop-up and pop-under advertisements that were encountered as the participants completed the tasks.

5.1.4   Data Distribution and Analysis Techniques

Many analysis techniques such as ANOVA rely on an assumption that the data being analyzed follows a normal distribution. In the data collected for this study, some of the dependent variables had skewed distributions, often to a ceiling or floor. ANOVA can be robust to some violation of the assumption of normality. In order to investigate this, each of the major analyses presented in this chapter was run once with the data as collected, and again with a log transformation applied to the data. Applying a log transform is a common technique used to adjust skewed data into a more normal distribution. A comparison of the two analyses – the original data and the log-transformed data – showed that almost all the statistical tests and effects were the same between the two analyses. A description of the log transform used and comparison of the log transform analyses to the analyses of the original data is presented in Appendix A.

Modern computer statistical software allows sophisticated analysis techniques to be used to analyze data. For the analysis presented here, SAS *proc mixed* was used to conduct the ANOVAs. Its features include the ability to handle designs with unbalanced and missing cells and good handling of repeated measures designs. Using *proc mixed* helped address the unbalanced and missing cells in the data and the repeated measures design.

5.2   Analysis 1 – Controlled Factors Analysis

This analysis examines the effects of controlled factors Session (D) and Task (T) on the measured variables: task familiarity, task frequency (see note about task frequency in section 4.4.4.1), after-task difficulty, and before-task difficulty, and on the dependent variables: task completion time, number of URLs, number of searches, and search ratio. Because this analysis compares data across sessions, only the EP-15 tasks are included (these are the exact (E) and path (P) tasks except task 15).

This analysis addresses Hypotheses 1 and 3:

> *Hypothesis 1 – Finding and Refinding have differences when defined by Session*
> *There will be a significant effect of Session on task completion time, number of URLs, number of searches, search ratio, familiarity, before-task difficulty, and after-task difficulty, but not on frequency. Specifically, task completion time, number of URLs, search ratio, before-task difficulty, and after-task difficulty will all be lower/easier for Session 2 than for Session 1. Familiarity will be higher for Session 2 than for Session 1.*

> *Hypothesis 3 – Task has a large effect on finding and refinding*
> *There will be a significant effect of Task on the dependent measures task completion time, number of URLs, number of searches, search ratio, familiarity, frequency, before-task difficulty, and after-task difficulty for the Session 1 and Session 2 tasks. Post-hoc tests will indicate the effects on a per task basis.*

Two-way within-subjects ANOVAs were run on the EP-15 tasks from both session 1 and session 2 to examine the effects of Session and Task on the dependent variables. Table 8 shows the ANOVA design.

Table 8. Within-Subjects ANOVA for Session x Task (EP-15 S12 tasks)

Source	*Numerator df*	*Denominator df*
Session (D)	1	16
Task (T)	11	151
DxT	11	135-138*

       * as reported by SAS proc mixed

The results of the ANOVAs are shown in Table 9. The interaction effect of Session x Task was significant at p<= 0.01 for task completion time, task familiarity, and before-task difficulty. The Session x Task interaction was also marginally significant at p=0.06 for number of URLs. The main effect of Task was significant at p<=0.01for all the measured and dependent variables. The main effect of Session was significant at p<=0.01 for task completion time, number or URLs, task familiarity, before-task difficulty, and after-task difficulty.

Table 9. Results of Two-Way ANOVA for Session x Task (EP-15 S12 tasks)

	Session (D)		*Task (T)*		*Interaction (DxT)*	
	F	*P*	*F*	*P*	*F*	*P*
Task Completion Time	20.06	<.01*	31.42	<.01*	2.61	<.01*
Number of URLs	9.10	0.01*	12.46	<.01*	1.78	0.06
Number of Searches	3.47	0.08	8.41	<.01*	0.64	0.80
Search Ratio	0.12	0.73	13.55	<.01*	0.20	1.00
Task Familiarity	49.80	<.01*	15.23	<.01*	3.60	<.01*
Task Frequency~	0.12	0.74	18.48	<.01*	0.53	0.88
Before-task Difficulty	37.88	<.01*	15.97	<.01*	3.29	<.01*
After-task Difficulty	21.98	<.01*	10.99	<.01*	1.23	0.28

* p ≤ 0.01, ~ see note about task frequency in section 4.4.4.1

5.2.1   Effects of Session x Task Interaction

The Session x Task interaction had a significant effect at p<=0.01 on task completion time, task familiarity, and before-task difficulty. Two additional analyses were conducted to gain a better understanding of these effects. First, post-hoc analyses using the Tukey-Kramer adjustment were conducted on the Session x Task interaction for the dependent variables task completion time, task familiarity, and before-task difficulty. These post-hoc analyses compared every combination of Session x Task with every other combination for each of the dependent variables. The results of the post-hoc analyses are shown in Figure 7, Figure 9, and Figure 11.

Second, follow-up one-way ANOVAs were run for each task with Session as the independent variable and task completion time, task familiarity, and before-task difficulty as the dependent variable. The follow-up ANOVAs give insight into how the dependent variables differed on a particular task from the first session to the second session[9]. The results of the follow-up ANOVAs are summarized in Table 10.

---

[9] The table of post-hoc comparisons could be used to examine the dependent variables across the two sessions, but I believe that conducting a follow-up one-way ANOVA for each task provides a more precise measure since the number of participants that completed each task was different for each task and is better accounted for in the individual ANOVAs in terms of comparing task performance across sessions.

Analyses of the Session x Task interaction are presented in the following pages. Each page contains the following results of the analysis of one of the dependent variables that was significantly affected by the Session x Task interaction:

- A chart of the mean values and standard error values for each EP-15 task for the dependent variable

- Stars ("*") on the chart to indicate which tasks were different from session 1 to session 2 based on the follow-up ANOVA analyses.

- A matrix of the post-hoc comparisons of the differences of least square means showing the Tukey-Kramer adjusted p-values comparing each task to every other task for each session.

Finally, on page 60, Table 11 shows the numeric mean values for each of the dependent variables for the tasks that were significantly different from session 1 to session 2 based on the follow-up one-way ANOVA tests.

Table 10. Results of One-Way ANOVA for Session on Each EP-15 task (EP-15 S12 tasks)

Task	Description	Task Completion Time p-value	Task Familiarity p-value	BeforeTask Difficulty p-value
1	Dictionary	0.51	0.10	0.06
3	Mish Mish phone	0.36	0.01 *	<0.01 *
5	Flight reservation	0.02 **	0.10	0.37
6	Rubik's cube	0.02 **	<0.01 *	<0.01 *
8	Kroger phone	0.09	<0.01 *	<0.01 *
10	Football score	0.71	0.01 *	0.07
11	Movie times	0.46	0.94	0.86
12	Stock price	0.12	0.04 **	0.01 *
13	Stasko Web site	0.04 **	<0.01 *	<0.01 *
16	Local temperatures	0.06	0.58	0.10
17	Academic calendar	0.75	0.67	0.22
18	Map distance	0.09	0.14	0.87

* p ≤ 0.01, ** p ≤ 0.05
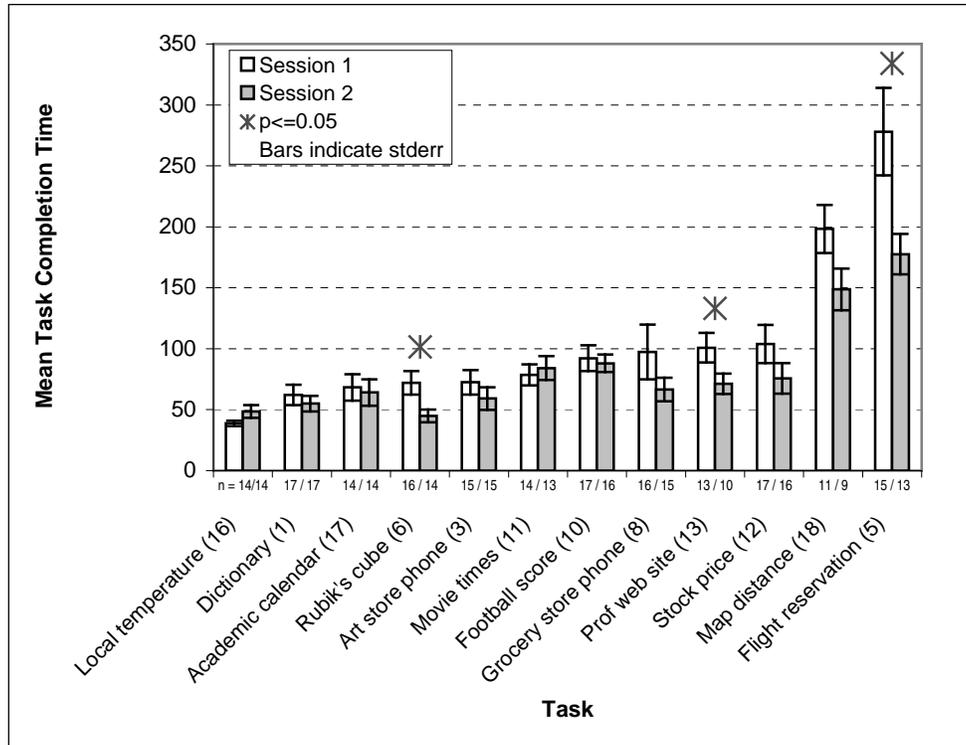
Figure 6. Mean Task Completion Time for Session x Task

Session 1 tasks compared with Session 1 tasks   cell values indicate p-values; 0.00 indicates a value less than 0.0001
shaded cells are values less than or equal to 0.01

	16	1	17	6	3	11	10	8	13	12	18	5
16	1.00	0.27	0.12	0.11	0.09	0.03	0.01	0.00	0.00	0.00	0.00	0.00
1		1.00	0.60	0.60	0.51	0.25	0.07	0.03	0.02	0.01	0.00	0.00
17			1.00	0.99	0.90	0.55	0.23	0.13	0.09	0.06	0.00	0.00
6				1.00	0.89	0.52	0.21	0.11	0.08	0.05	0.00	0.00
3					1.00	0.62	0.28	0.15	0.11	0.08	0.00	0.00
11						1.00	0.58	0.36	0.27	0.22	0.00	0.00
10							1.00	0.71	0.55	0.48	0.00	0.00
8								1.00	0.81	0.75	0.00	0.00
13									1.00	0.96	0.00	0.00
12										1.00	0.00	0.00
18											1.00	0.00
5												1.00

Session 1 tasks (y-axis) compared with Session 2 tasks (x-axis)

	16	1	17	6	3	11	10	8	13	12	18	5
16	0.59	0.49	0.19	0.98	0.34	0.02	0.01	0.16	0.18	0.07	0.00	0.00
1	0.59	0.67	0.77	0.29	0.91	0.17	0.13	0.72	0.67	0.43	0.00	0.00
17	0.31	0.35	0.82	0.13	0.54	0.41	0.35	0.87	0.96	0.82	0.00	0.00
6	0.30	0.34	0.83	0.12	0.53	0.39	0.33	0.88	0.97	0.80	0.00	0.00
3	0.25	0.28	0.73	0.10	0.45	0.47	0.41	0.77	0.87	0.91	0.00	0.00
11	0.11	0.12	0.41	0.04	0.22	0.82	0.76	0.44	0.55	0.69	0.00	0.00
10	0.03	0.03	0.15	0.01	0.06	0.76	0.80	0.17	0.26	0.32	0.00	0.00
8	0.01	0.01	0.08	0.00	0.03	0.51	0.54	0.08	0.15	0.18	0.01	0.00
13	0.01	0.01	0.06	0.00	0.02	0.40	0.41	0.06	0.11	0.13	0.02	0.00
12	0.00	0.00	0.04	0.00	0.01	0.34	0.35	0.04	0.09	0.09	0.01	0.00
18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.17
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Session 2 tasks compared with Session 2 tasks

	16	1	17	6	3	11	10	8	13	12	18	5
16	1.00	0.90	0.43	0.61	0.68	0.07	0.05	0.39	0.39	0.20	0.00	0.00
1		1.00	0.49	0.51	0.77	0.08	0.05	0.44	0.43	0.23	0.00	0.00
17			1.00	0.20	0.70	0.30	0.24	0.95	0.88	0.64	0.00	0.00
6				1.00	0.35	0.02	0.01	0.17	0.18	0.07	0.00	0.00
3					1.00	0.15	0.11	0.64	0.61	0.38	0.00	0.00
11						1.00	0.95	0.32	0.42	0.53	0.00	0.00
10							1.00	0.26	0.37	0.47	0.00	0.00
8								1.00	0.92	0.69	0.00	0.00
13									1.00	0.79	0.00	0.00
12										1.00	0.00	0.00
18											1.00	0.30
5												1.00

Figure 7. Post-hoc comparisons of Mean Task Completion Time for Session x Task
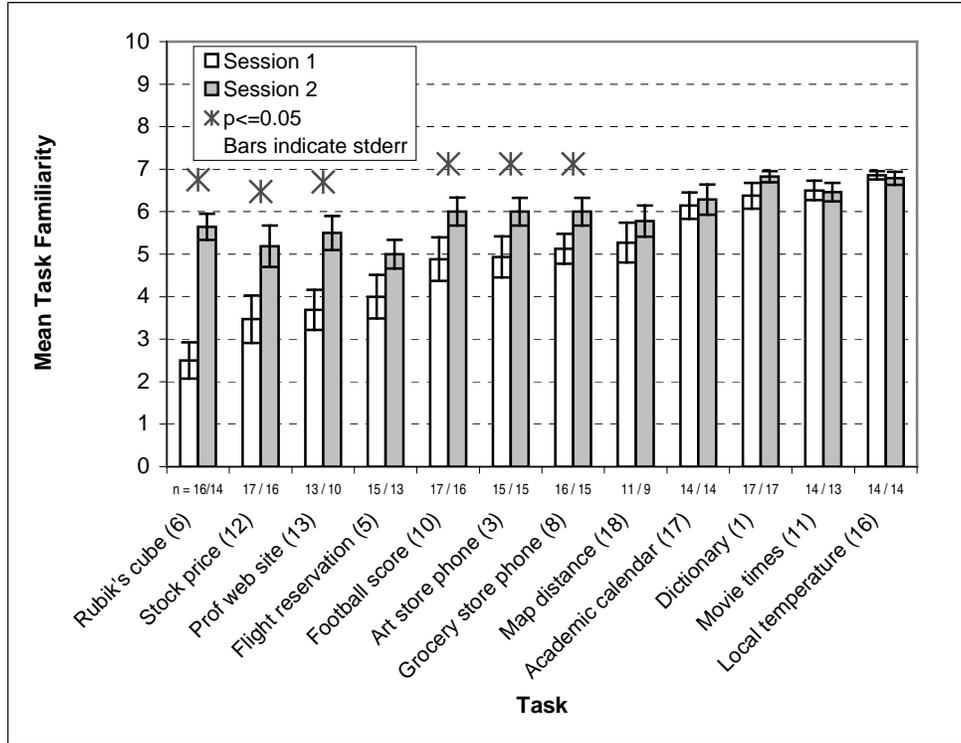
Figure 8. Mean Task Familiarity for Session x Task

Session 1 tasks compared with Session 1 tasks — cell values indicate p-values; 0.00 indicates a value less than 0.0001; shaded cells are values less than or equal to 0.01

	6	12	13	5	10	3	8	18	17	1	11	16
6	1.00	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12		1.00	0.70	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13			1.00	0.36	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00
5				1.00	0.09	0.09	0.03	0.06	0.00	0.00	0.00	0.00
10					1.00	0.92	0.61	0.68	0.01	0.00	0.00	0.00
3						1.00	0.69	0.75	0.01	0.00	0.00	0.00
8							1.00	0.96	0.03	0.01	0.00	0.00
18								1.00	0.05	0.01	0.01	0.00
17									1.00	0.59	0.48	0.15
1										1.00	0.85	0.35
11											1.00	0.47
16												1.00

Session 1 tasks (y-axis) compared with Session 2 tasks (x-axis)

	6	12	13	5	10	3	8	18	17	1	11	16
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.02	0.02	0.03	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
10	0.13	0.44	0.33	0.55	0.01	0.02	0.02	0.17	0.00	0.00	0.00	0.00
3	0.17	0.52	0.38	0.63	0.02	0.02	0.02	0.20	0.01	0.00	0.00	0.00
8	0.32	0.80	0.60	0.91	0.04	0.06	0.06	0.34	0.01	0.00	0.00	0.00
18	0.34	0.78	0.60	0.89	0.06	0.08	0.08	0.36	0.02	0.00	0.01	0.00
17	0.26	0.06	0.17	0.05	0.85	0.78	0.78	0.36	0.77	0.13	0.45	0.20
1	0.09	0.01	0.06	0.01	0.45	0.40	0.40	0.16	0.81	0.32	0.81	0.43
11	0.07	0.01	0.04	0.01	0.36	0.32	0.32	0.12	0.68	0.44	0.96	0.56
16	0.01	0.00	0.01	0.00	0.10	0.08	0.08	0.03	0.25	1.00	0.51	0.88

Session 2 tasks compared with Session 2 tasks

	6	12	13	5	10	3	8	18	17	1	11	16
6	1.00	0.45	0.71	0.40	0.33	0.39	0.39	0.95	0.16	0.01	0.07	0.02
12		1.00	0.76	0.90	0.07	0.10	0.10	0.46	0.03	0.00	0.01	0.00
13			1.00	0.68	0.21	0.25	0.25	0.69	0.10	0.01	0.04	0.01
5				1.00	0.07	0.09	0.09	0.42	0.03	0.00	0.01	0.00
10					1.00	0.92	0.92	0.44	0.62	0.08	0.34	0.13
3						1.00	1.00	0.49	0.57	0.07	0.30	0.11
8							1.00	0.49	0.57	0.07	0.30	0.11
18								1.00	0.24	0.02	0.12	0.04
17									1.00	0.23	0.64	0.32
1										1.00	0.49	0.88
11											1.00	0.61
16												1.00

Figure 9. Post-hoc comparison of Mean Task Familiarity for Session x Task
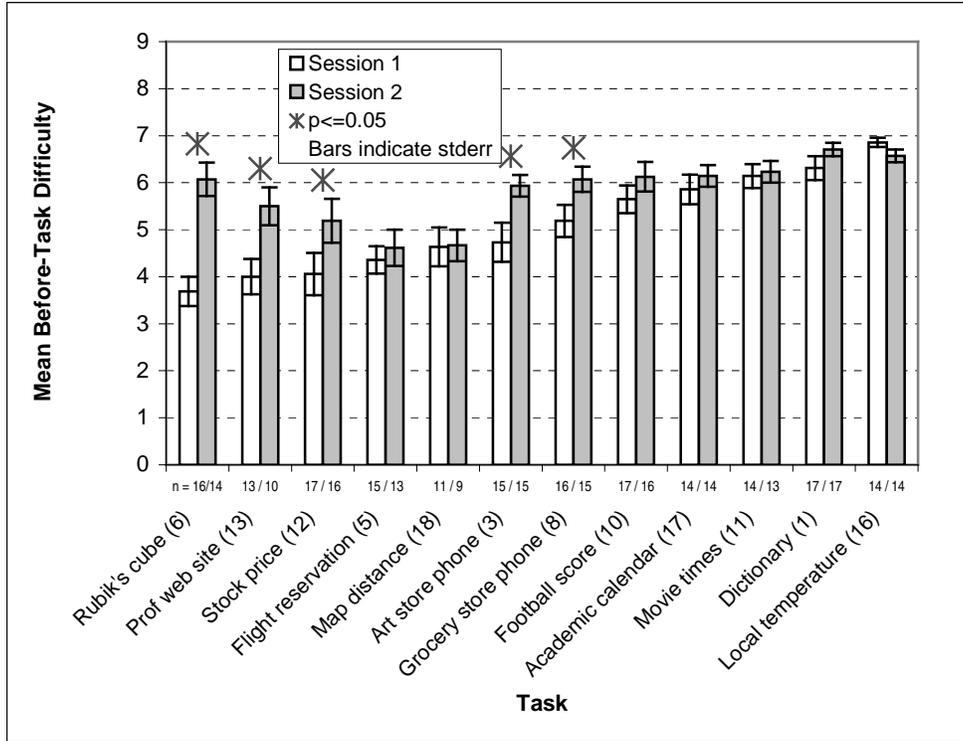
Figure 10. Mean Before-Task Difficulty for Session x Task

**Session 1 tasks compared with Session 1 tasks** — cell values indicate p-values; 0.00 indicates a value less than 0.0001; shaded cells are values less than or equal to 0.01

	6	13	12	5	18	3	8	10	17	11	1	16
6	1.00	0.42	0.27	0.05	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00
13		1.00	0.82	0.28	0.17	0.07	0.00	0.00	0.00	0.00	0.00	0.00
12			1.00	0.36	0.21	0.08	0.00	0.00	0.00	0.00	0.00	0.00
5				1.00	0.71	0.45	0.04	0.00	0.00	0.00	0.00	0.00
18					1.00	0.74	0.12	0.01	0.00	0.00	0.00	0.00
3						1.00	0.19	0.01	0.00	0.00	0.00	0.00
8							1.00	0.23	0.11	0.01	0.00	0.00
10								1.00	0.64	0.15	0.09	0.00
17									1.00	0.36	0.24	0.01
11										1.00	0.83	0.10
1											1.00	0.14
16												1.00

Session 1 tasks (y-axis) compared with Session 2 tasks (x-axis)

	6	12	13	5	10	3	8	18	17	1	11	16
6	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.06	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.07	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.05	0.04	0.38	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.13	0.13	0.65	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.20	0.19	0.88	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.04	0.89	0.99	0.27	0.24	0.06	0.01	0.01	0.02	0.00	0.00	0.00
18	0.35	0.36	0.22	0.02	0.03	0.48	0.18	0.18	0.22	0.08	0.00	0.02
17	0.66	0.20	0.11	0.01	0.01	0.83	0.40	0.41	0.47	0.22	0.02	0.06
1	0.63	0.04	0.01	0.00	0.00	0.47	0.94	0.89	0.84	0.74	0.18	0.35
11	0.48	0.02	0.00	0.00	0.00	0.33	0.77	0.72	0.67	0.89	0.24	0.45
16	0.03	0.00	0.00	0.00	0.00	0.02	0.09	0.07	0.06	0.20	0.71	0.47

Session 2 tasks compared with Session 2 tasks

	6	12	13	5	10	3	8	18	17	1	11	16
6	1.00	0.09	0.04	0.00	0.00	0.81	0.69	0.72	0.78	0.42	0.06	0.16
12		1.00	0.88	0.27	0.24	0.14	0.04	0.04	0.05	0.02	0.00	0.00
13			1.00	0.27	0.25	0.06	0.01	0.01	0.02	0.00	0.00	0.00
5				1.00	0.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10					1.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
3						1.00	0.52	0.54	0.61	0.30	0.03	0.10
8							1.00	0.96	0.90	0.68	0.15	0.31
18								1.00	0.94	0.63	0.12	0.27
17									1.00	0.59	0.12	0.26
1										1.00	0.33	0.56
11											1.00	0.70
16												1.00

Figure 11. Post-hoc comparisons of Mean Before-Task Difficulty for Session x Task

Table 11. Session x Task – Time and Number of URLs Means (EP-15 S12 tasks)

Task	Description	Task Completion Time		Task Familiarity		Before Task Difficulty	
		S1	*S2*	*S1*	*S2*	*S1*	*S2*
3	*Mish Mish phone*	--	--	4.93	6.00	4.73	5.93
5	*Flight reservation*	278.07	177.54	--	--	--	--
6	*Rubik's cube*	72.00	44.86	2.50	5.64	3.69	6.07
8	*Kroger phone*	--	--	5.13	6.00	5.19	6.07
10	*Football score*	--	--	4.88	6.00	--	--
12	*Stock price*	--	--	3.47	5.19	4.06	5.19
13	*Stasko Web site*	103.88	71.20	3.69	5.50	4.00	5.50
16	*Local temperatures*	--	--	--	--	--	--

* cells with a "--" are blank because that variable was not significant in the Session x Task interaction

5.2.2   Effects of Task

Task had a significant main effect at $p \le 0.01$ on all the dependent variables (task completion time, number or URLs, number of searches, search ratio, task familiarity, task frequency (see note about task frequency in section 4.4.4.1), before-task difficulty, and after-task difficulty). Post-hoc analyses using the Tukey-Kramer adjustment were conducted for each of the dependent variables to examine the significant differences of Task.

Analyses of the main effect of Task are presented in the following pages. Each page contains the following results of the analysis of one of the dependent variables that were significantly affected by Task:

- A chart of the mean values and standard error values for each EP-15 task for the dependent variable

- A matrix of the post-hoc comparisons of the differences of least square means showing the Tukey-Kramer adjusted p-values comparing each task to every other task.

5.2.2.1   Task Completion Time
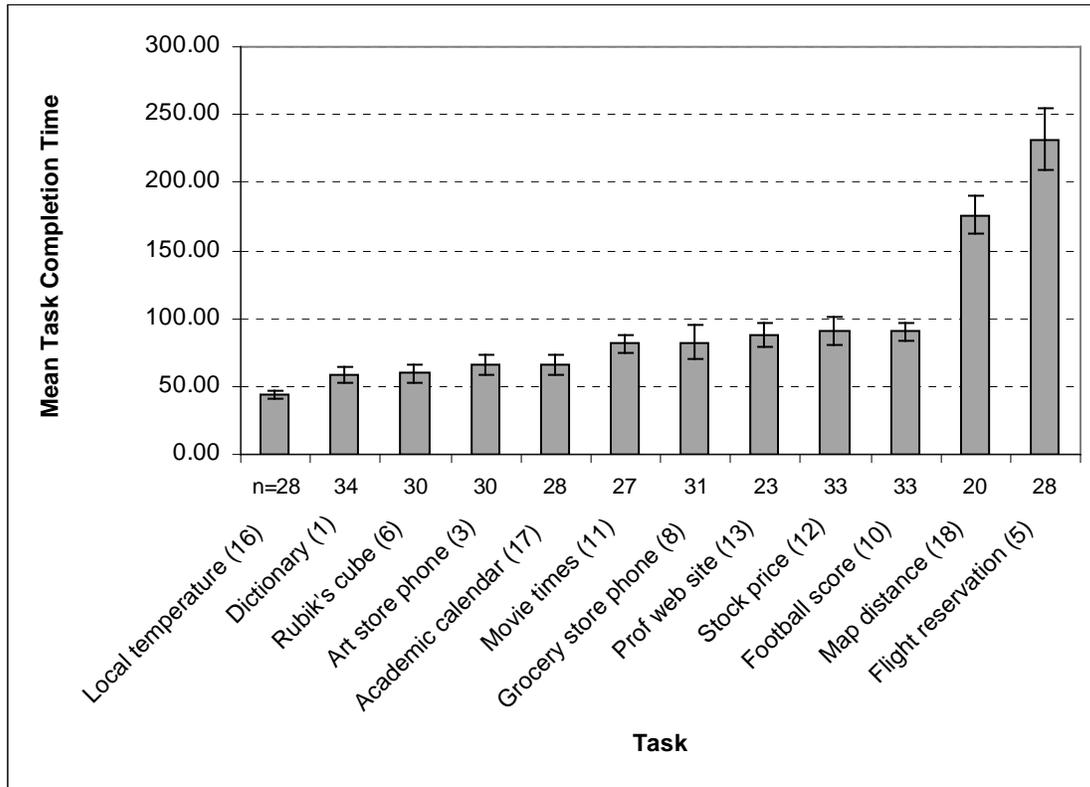
Figure 12. Graph of Task effect on Mean Task Completion Time (EP-15 S12 tasks)

Task Completion Time -- Post-hoc comparisons
cells indicate p-values; 0.00 indicates a value less than 0.01

Task	16	1	6	3	17	11	8	13	12	10	18	5
16	1.00	0.39	0.47	0.14	0.10	0.01	0.01	0.01	0.00	0.00	0.00	0.00
1		1.00	0.91	0.50	0.39	0.04	0.04	0.04	0.01	0.01	0.00	0.00
6			1.00	0.45	0.35	0.04	0.04	0.03	0.01	0.01	0.00	0.00
3				1.00	0.85	0.17	0.18	0.15	0.06	0.06	0.00	0.00
17					1.00	0.25	0.26	0.21	0.10	0.10	0.00	0.00
11						1.00	0.93	0.88	0.69	0.66	0.00	0.00
8							1.00	0.81	0.61	0.59	0.00	0.00
13								1.00	0.82	0.80	0.00	0.00
12									1.00	0.97	0.00	0.00
10										1.00	0.00	0.00
18											1.00	0.00
5												1.00

Groups

	16	1	6	3	17	11	8	13	12	10	18	5
	A	A	A	A	A							
	B	B	B	B	B	B	B					
	C	C	C	C	C	C	C	C	C			
											D	
												E

Figure 13. Post-hoc task comparisons of Task effect on Task Completion Time (EP-15 S12 tasks)

5.2.2.2   Number of URLs
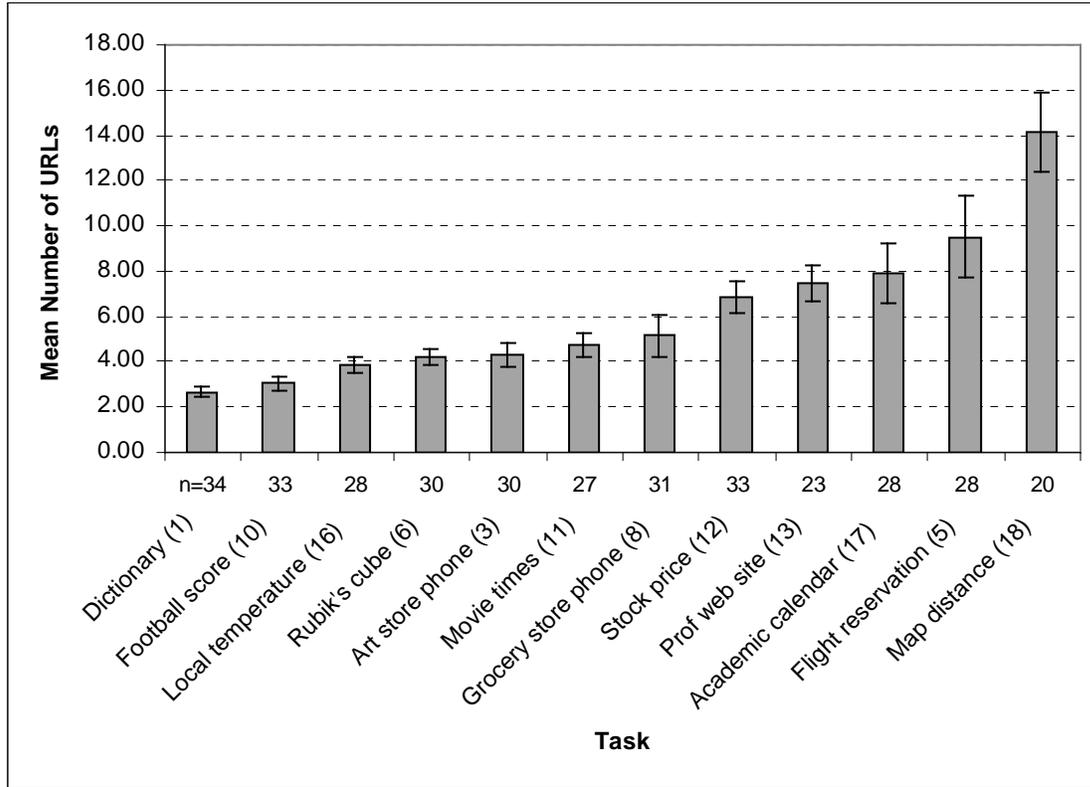
Figure 14. Graph of Task effect on Mean Number of URLs (EP-15 S12 tasks)

Number of URLs -- Post-hoc comparisons
cells indicate p-values; 0.00 indicates a value less than 0.01

Task	1	10	16	6	3	11	8	12	13	17	5	18
1	1.00	0.78	0.18	0.19	0.15	0.06	0.03	0.00	0.00	0.00	0.00	0.00
10		1.00	0.29	0.30	0.25	0.10	0.05	0.00	0.00	0.00	0.00	0.00
16			1.00	0.97	0.94	0.57	0.40	0.02	0.02	0.00	0.00	0.00
6				1.00	0.91	0.54	0.38	0.02	0.02	0.00	0.00	0.00
3					1.00	0.61	0.44	0.03	0.02	0.00	0.00	0.00
11						1.00	0.81	0.10	0.08	0.01	0.00	0.00
8							1.00	0.14	0.11	0.02	0.00	0.00
12								1.00	0.79	0.30	0.05	0.00
13									1.00	0.50	0.12	0.00
17										1.00	0.35	0.00
5											1.00	0.00
18												1.00

Groups

	1	10	16	6	3	11	8	12	13	17	5	18
	A	A	A	A	A	A	A					
			B	B	B	B	B	B	B			
						C	C	C	C	C		
								D	D	D	D	
												E

Figure 15. Post-hoc task comparisons of Task effect on Mean Number of URLs (EP-15 S12 tasks)

5.2.2.3   Number of searches
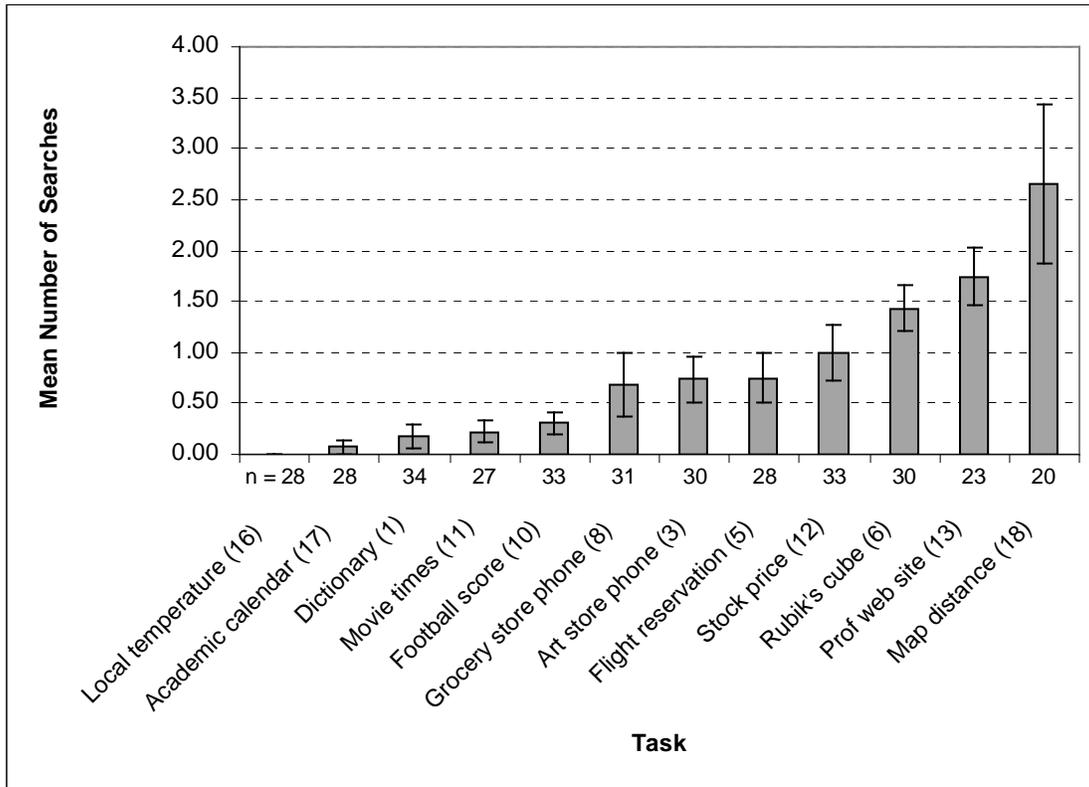
Figure 16. Graph of Task effect on Mean Number of Searches (EP-15 S12 tasks)

Number of Searches -- Post-hoc comparisons
cells indicate p-values; 0.00 indicates a value less than 0.01

Task	16	17	1	11	10	8	3	5	12	6	13	18
16	1.00	0.92	0.86	0.46	0.58	0.10	0.08	0.10	0.01	0.00	0.00	0.00
17		1.00	0.77	0.40	0.51	0.08	0.07	0.08	0.01	0.00	0.00	0.00
1			1.00	0.55	0.70	0.12	0.10	0.12	0.01	0.00	0.00	0.00
11				1.00	0.82	0.37	0.33	0.37	0.06	0.00	0.00	0.00
10					1.00	0.24	0.21	0.24	0.03	0.00	0.00	0.00
8						1.00	0.93	0.98	0.32	0.02	0.01	0.00
3							1.00	0.95	0.37	0.03	0.01	0.00
5								1.00	0.35	0.03	0.01	0.00
12									1.00	0.18	0.06	0.00
6										1.00	0.51	0.00
13											1.00	0.01
18												1.00

Groups

	16	17	1	11	10	8	3	5	12	6	13	18
	A	A	A	A	A	A	A	A				
				B	B	B	B	B	B			
						C	C	C	C	C		
									D	D	D	
											E	E

Figure 17. Post-hoc task comparisons of Task effect on Number of Searches (EP-15 S12 tasks)

5.2.2.4   Search Ratio
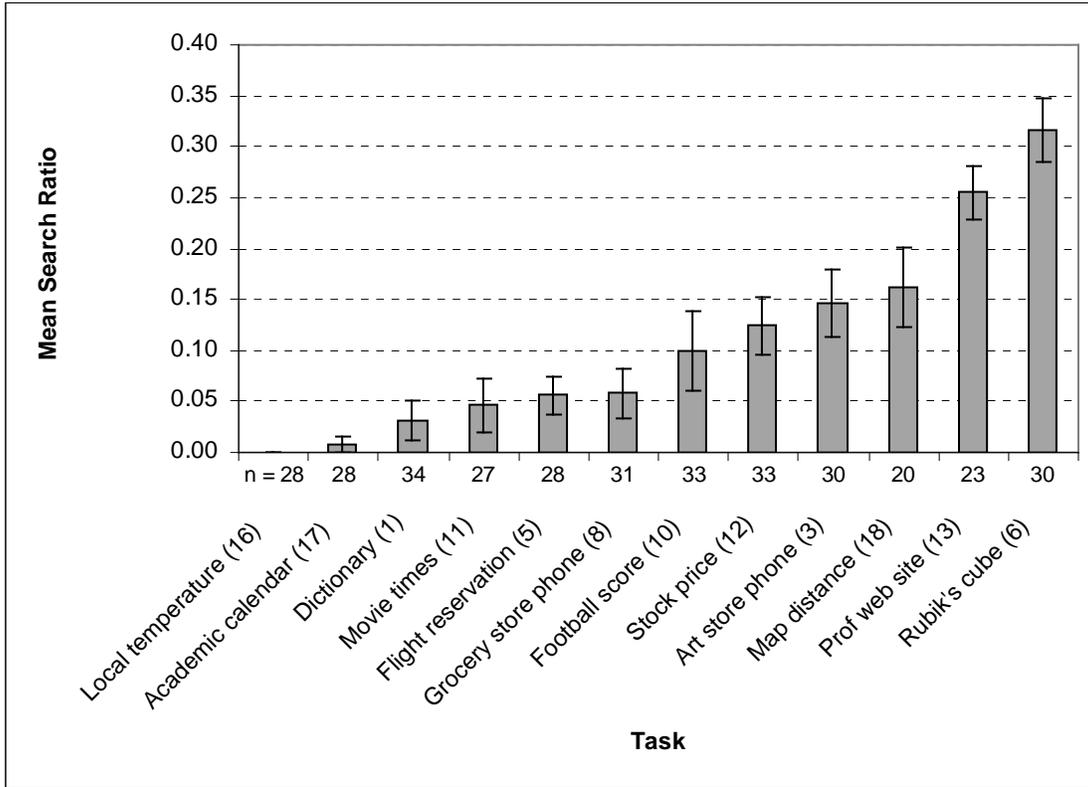
Figure 18. Graph of Task effect on Mean Search Ratio (EP-15 S12 tasks)

Search Ratio -- Post-hoc comparisons
cells indicate p-values; 0.00 indicates a value less than 0.01

Task	16	17	1	5	11	8	10	12	3	18	13	6
16	1.00	0.99	0.46	0.22	0.13	0.14	0.01	0.00	0.00	0.00	0.00	0.00
17		1.00	0.46	0.22	0.13	0.14	0.01	0.00	0.00	0.00	0.00	0.00
1			1.00	0.58	0.39	0.43	0.05	0.01	0.00	0.00	0.00	0.00
5				1.00	0.76	0.83	0.18	0.04	0.01	0.01	0.00	0.00
11					1.00	0.93	0.31	0.09	0.03	0.02	0.00	0.00
8						1.00	0.25	0.06	0.02	0.01	0.00	0.00
10							1.00	0.47	0.22	0.12	0.00	0.00
12								1.00	0.60	0.34	0.00	0.00
3									1.00	0.63	0.01	0.00
18										1.00	0.04	0.00
13											1.00	0.09
6												1.00

Groups

A	A	A	A	A	A						
		B	B	B	B	B					
		C	C	C	C	C	C				
			D	D	D	D	D				
								D	E	E	
											F

Figure 19. Post-hoc task comparisons of Task effect on Search Ratio (EP-15 S12 tasks)

5.2.2.5   Task Frequency
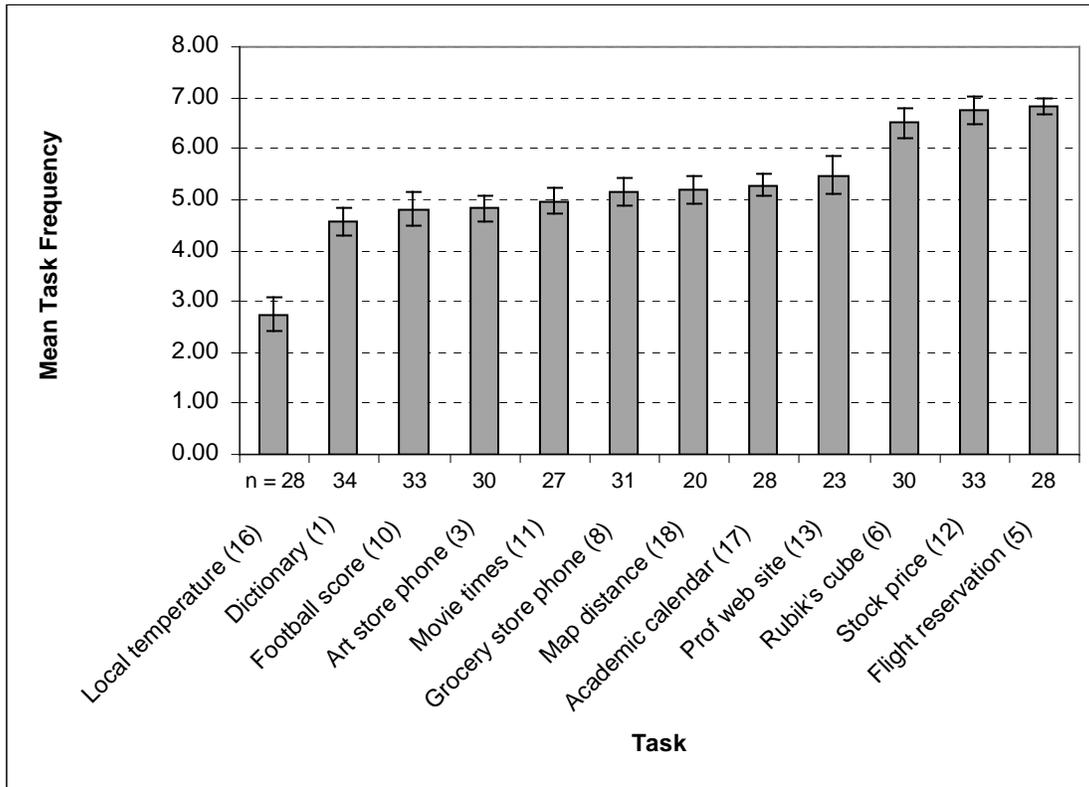
Figure 20. Graph of Task effect on Mean Task Frequency (EP-15 S12 tasks)

No post-hoc comparison table was generated for the task frequency data (see note about task frequency in section 4.4.4.1)

5.2.2.6   *Task Familiarity*
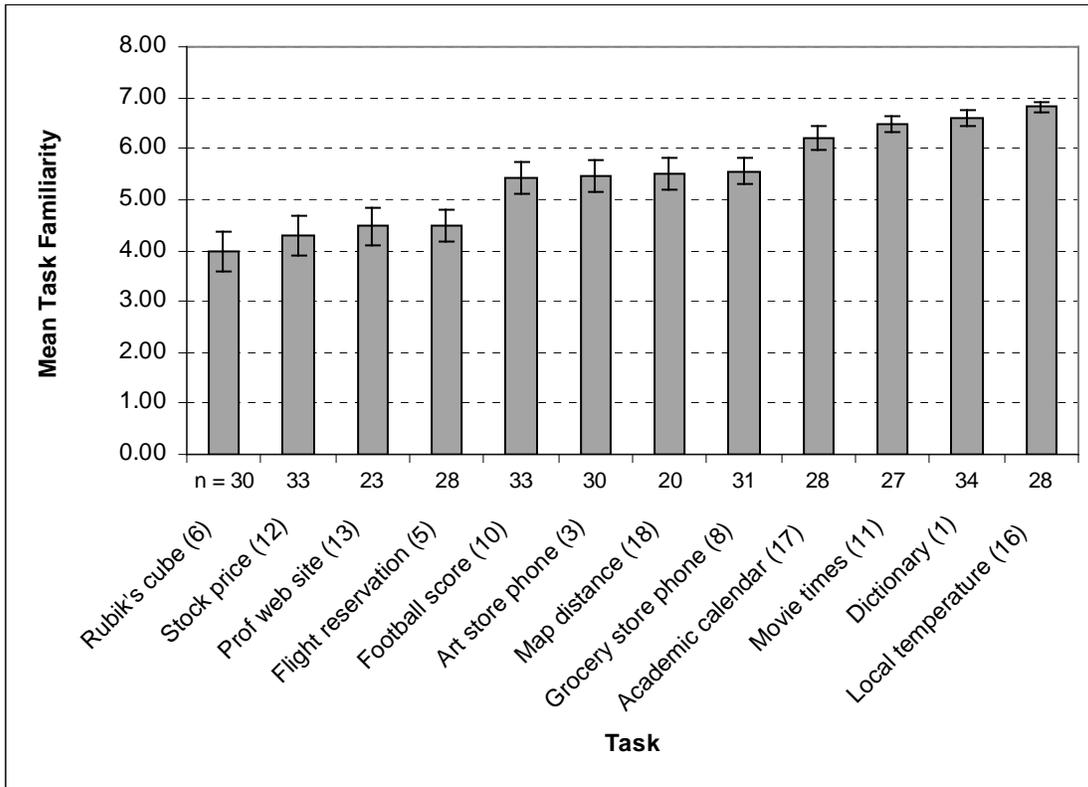
Figure 21. Graph of Task effect on Mean Task Familiarity (EP-15 S12 tasks)

Task Familiarity -- Post-hoc comparisons
cells indicate p-values; 0.00 indicates a value less than 0.01

	6	12	13	5	18	10	3	8	17	11	1	16
6	1.00	0.32	0.17	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12		1.00	0.62	0.39	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13			1.00	0.76	0.04	0.01	0.01	0.00	0.00	0.00	0.00	0.00
5				1.00	0.06	0.01	0.02	0.01	0.00	0.00	0.00	0.00
18					1.00	0.77	0.77	0.59	0.03	0.00	0.00	0.00
10						1.00	1.00	0.77	0.03	0.00	0.00	0.00
3							1.00	0.78	0.03	0.00	0.00	0.00
8								1.00	0.06	0.01	0.00	0.00
17									1.00	0.41	0.22	0.09
11										1.00	0.72	0.38
1											1.00	0.58
16												1.00

Groups

6	12	13	5	18	10	3	8	17	11	1	16
A	A	A	A								
	B	B	B								
		C	C	C	C						
			D	D	D	D		D			
							E	E	E	E	E

Figure 22. Post-hoc task comparisons of Task effect on Task Familiarity (EP-15 S12 tasks)

5.2.2.7   Before-Task Difficulty
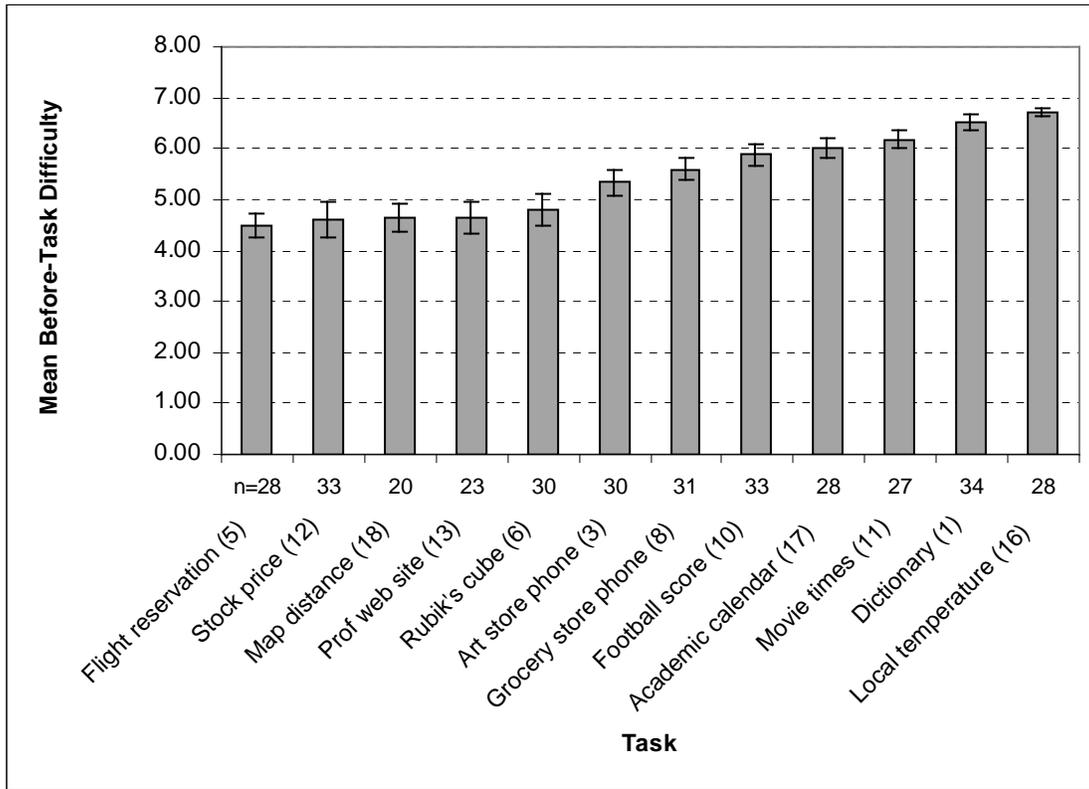
Figure 23. Graph of Task effect on Mean Before-Task Difficulty (EP-15 S12 tasks)

Before-Task Difficulty -- Post-hoc comparisons
cells indicate p-values; 0.00 indicates a value less than 0.01

Task	5	12	18	13	6	3	8	10	17	11	1	16
5	1.00	0.88	0.90	0.92	0.40	0.01	0.00	0.00	0.00	0.00	0.00	0.00
12		1.00	1.00	0.97	0.46	0.01	0.00	0.00	0.00	0.00	0.00	0.00
18			1.00	0.97	0.52	0.03	0.00	0.00	0.00	0.00	0.00	0.00
13				1.00	0.48	0.02	0.00	0.00	0.00	0.00	0.00	0.00
6					1.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00
3						1.00	0.17	0.03	0.02	0.00	0.00	0.00
8							1.00	0.43	0.31	0.04	0.00	0.00
10								1.00	0.79	0.18	0.02	0.00
17									1.00	0.31	0.05	0.01
11										1.00	0.40	0.12
1											1.00	0.43
16												1.00

Groups

	5	12	18	13	6	3	8	10	17	11	1	16
	A	A	A	A	A	A						
						B	B	B	B			
						C	C	C	C			
						D		D	D	D		
										E	E	E

Figure 24. Post-hoc task comparisons of Task effect on Before-Task Difficulty (EP-15 S12 tasks)

*5.2.2.8   After-Task Difficulty*
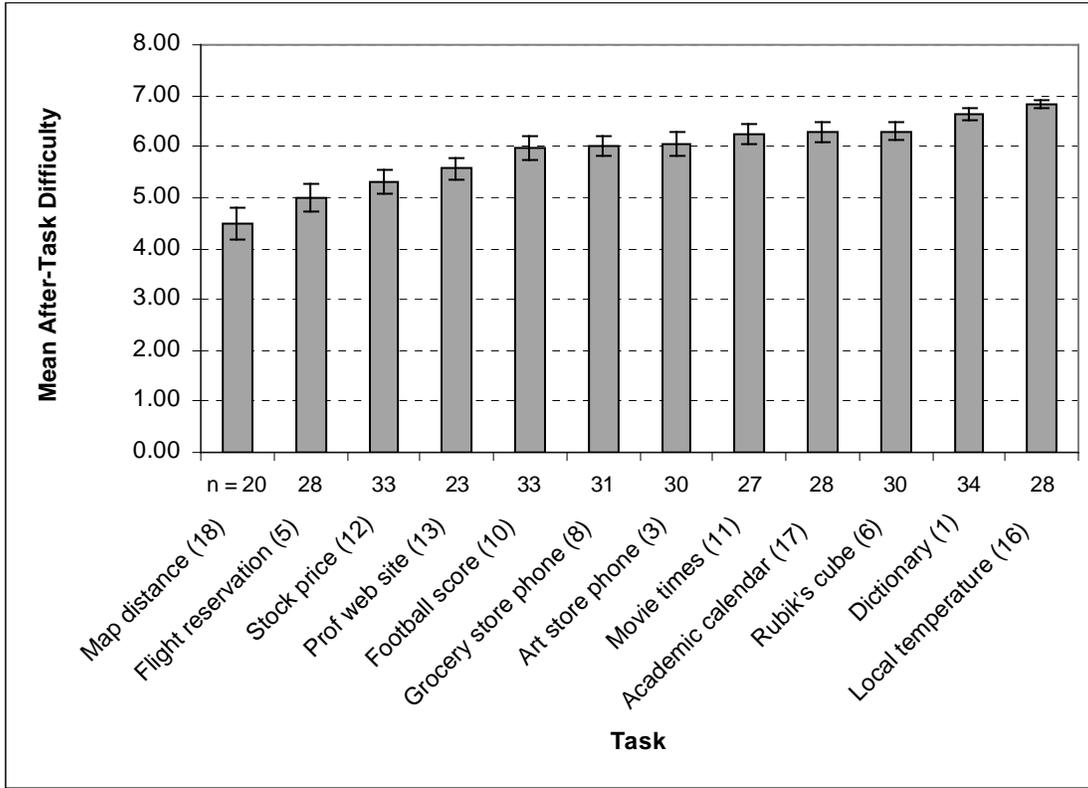
Figure 25. Graph of Task effect on Mean After-Task Difficulty (EP-15 S12 tasks)

After-Task Difficulty -- Post-hoc comparisons
cells indicate p-values; 0.00 indicates a value less than 0.01

Task	18	5	12	13	10	8	3	11	17	6	1	16
18	1.00	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5		1.00	0.48	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12			1.00	0.50	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13				1.00	0.09	0.05	0.05	0.01	0.01	0.01	0.00	0.00
10					1.00	0.76	0.73	0.18	0.27	0.25	0.01	0.00
8						1.00	0.97	0.31	0.43	0.40	0.02	0.00
3							1.00	0.33	0.46	0.43	0.02	0.00
11								1.00	0.81	0.84	0.20	0.07
17									1.00	0.97	0.12	0.04
6										1.00	0.13	0.04
1											1.00	0.50
16												1.00

Groups

	18	5	12	13	10	8	3	11	17	6	1	16
	A	A										
		B	B	B								
				C	C	C	C					
					D	D	D	D	D	D		
					E		E	E	E	E	E	
								F	F	F	F	F

Figure 26. Post-hoc task comparisons of Task effect on After-Task Difficulty (EP-15 S12 tasks)

5.2.3   Effects of Session

Session had a significant main effect at $p \leq 0.01$ on task completion time, number of URLs, task familiarity, before-task difficulty and after-task difficulty. No effect of Session was observed for number of searches, search ratio, or task frequency (see note about task frequency in section 4.4.4.1). Table 12 shows the means and standard error values for each of the dependent variables that were significantly affected by Session. Figure 27 shows graphs of the means and standard error values for each of the dependent variables.

Table 12. Session Mean and Standard Error Values

	Session 1		Session 2		p-value
	Mean	*StdErr*	*Mean*	*StdErr*	
Task Completion Time	103.07	6.52	78.89	4.00	<.01
Number of URLs	6.51	0.47	5.10	0.34	0.01
Task Familiarity	4.94	0.15	5.98	0.10	<.01
Before-task Difficulty	5.12	0.12	5.88	0.10	<.01
After-task Difficulty	5.70	0.10	6.21	0.08	<.01
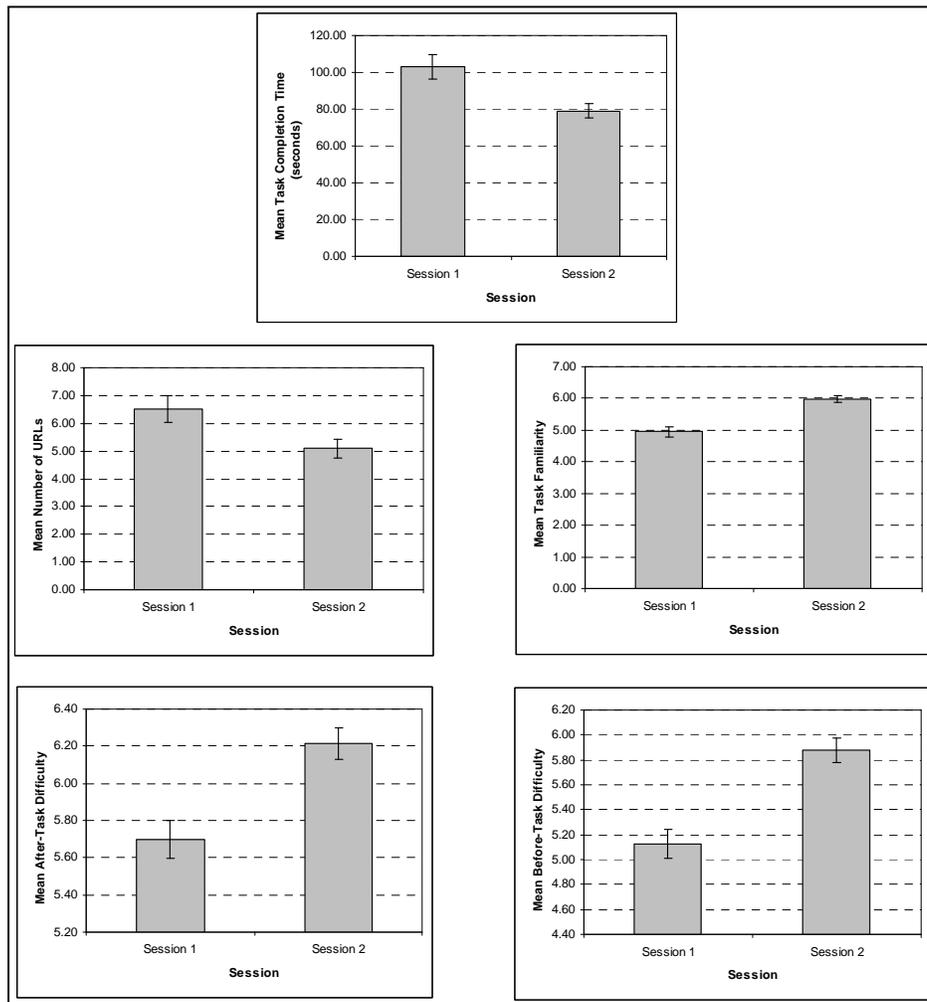
**Figure 27. Session Effect Mean Values**

5.2.4   Summary

In this section, the analysis is summarized in terms of the hypotheses for this analysis.

> *Hypothesis 1 – Finding and Refinding have differences when defined by Session*
> *There will be a significant effect of Session on task completion time, number of URLs, number of searches, search ratio, familiarity, before-task difficulty, and after-task difficulty, but not on frequency. Specifically, task completion time, number of URLs, search ratio, before-task difficulty, and after-task difficulty will all be lower/easier for Session 2 than for Session 1. Familiarity will be higher for Session 2 than for Session 1.*

Session had a significant main effect on task completion time, number of URLs, task familiarity, before-task difficulty, and after-task difficulty at $p<=0.01$. However, there was also a significant Session x Task interaction for task completion time, task familiarity, and before-task difficulty at $p<=0.01$.

Session x Task interaction – For the Session x Task interaction, it is the comparisons of a task in Session 1 to the same task in Session 2 that is of primary interest. In other words, it is interesting to consider the effect of session on each task, but not as interesting to compare task 1 in session 1 to task 2 in session 2. Post-hoc tests showed that task 6 (Rubik's cube) and task 13 (professor Web site) had significant differences ($p<=0.05$) based on session for task completion time, task familiarity, and before-task difficulty. Task 3 (Mish Mish phone), task 8 (Kroger phone), and task 12 (stock price) had significant differences based on session for task familiarity and before-task difficulty. Two other tasks had significant differences based on session for only one measure and are not further examined here (see Table 10 for details). When significant differences were found, task completion time was lower in session 2 than session 1, task familiarity in session 2 was higher in session 2 than session 1, and before-task difficulty was easier in session 2 than in session 1.

Session main effect – For the main effect of Session, there was a significant effect on task completion time, number of URLs, task familiarity, before-task difficulty, and after-task difficulty. Interestingly, no significant effect of Session was observed for number of searches or search ratio. As expected, no significant effect of Session was observed for task frequency (see note about task frequency in section 4.4.4.1). For the cases of significant differences in the mean values, task completion time was lower in session 2 than session 1, number of URLs was fewer in session 2 than session 1, task familiarity was higher in session 2 than session 1, before-task difficulty was easier in session 2 than in session 1, and after-task difficulty was easier in session 2 than in session 1.

Conclusion – the hypothesis is mainly supported by the analyses with several interesting additional observations. First, no main effect of session was observed for number of searches and search ratio. Second, although there was a main effect of Session, there were effects of the Session x Task interaction for several dependent measures. Analysis of these effects showed that some tasks were affected by session while for other tasks, no effect was observed. Tasks 6 (Rubik's cube), 13 (professor Web site), 3 (art store phone), 8 (grocery store phone), and 12 (stock price) were affected by the Session x Task interaction for two or more dependent measures.

> *Hypothesis 3 – Task has a large effect on finding and refinding*
> *There will be a significant effect of Task on the dependent measures task completion time, number of URLs, number of searches, search ratio, familiarity, frequency, before-task difficulty, and after-task difficulty for the Session 1 and Session 2 tasks. Post-hoc tests will indicate the effects on a per task basis.*

Effects of Task – Task had a significant main effect at $p<=0.01$ on all the dependent variables: task completion time, number or URLs, number of searches, search ratio, task familiarity, task frequency (see note about task frequency in section 4.4.4.1), before-task difficulty, and after-task difficulty. However, as

described in the summary for Hypothesis 1 above, there was also a significant Session x Task interaction for task completion time, task familiarity, and before-task difficulty at $p<=0.01$.

Post-hoc tests revealed differences based on task. For task completion time, task 18 (map distance) and task 5 (flight reservation) took significantly longer than all other tasks. The mean number of URLs was significantly greater for task 18 (map distance) than all other tasks. Task 6 (Rubik's cube) and task 13 (professor Web site) had significantly higher search ratios than all other tasks ($p<=0.05$). No tasks had higher familiarity than all the others, but three had higher familiarity than most: task 11 (movie times), task 1 (dictionary), and task 16 (local temperature). Similarly, task 16 (local temperature) was perceived as easier than most other tasks in both before-task and after-task difficulty measures. In the after-task difficulty measurement, task 18 (map distance) was perceived as significantly more difficult than all other tasks except one.

Conclusion – the hypothesis is mainly supported with some additional observations. Task did have a significant main effect on all the dependent variables. However, the interaction of Session x Tasks also had an effect on some tasks. Some tasks stood out in terms of their differences and not all tasks differed from all others – many tasks did not differ significantly in terms of the dependent measures even though they were in different topic areas. Most of the tasks in this study were fairly directed tasks that required finding a single item.

5.3   Analysis 2 – Prior Frequency Analysis

The analyses in this section consider finding and refinding based on the prior frequency with which participants did each task. Prior frequency is the frequency for each task as measured *before* the participant performed the task during Session 1. Prior frequency was mapped to three levels: low (less often than once per month), medium (once or several times per month), and high (once per week or more). These analyses examine the effects of factors session (D) and prior frequency (Q) on the dependent variables. Data from both sessions are included in this analysis. Note that prior frequency is the task frequency as measured during the first session, before the participant had performed the task. Because this analysis compares data across sessions, only the EP-15 tasks are included (these are the exact (E) and path (P) tasks except for task 15).

This analysis addresses Hypothesis 2:

> *Hypothesis 2 – Finding and Refinding have differences when defined by Task Frequency*
> *There will be a significant effect of task frequency on task completion time, number of URLs,*
> *number of searches, search ratio, familiarity, before-task difficulty, and after-task difficulty for the*
> *Session 1 and Session 2 tasks. Specifically, task completion time, number of URLs, number of*
> *searches, search ratio, before-task difficulty, and after-task difficulty will all be lower/easier for*
> *tasks with a high level of frequency than for tasks with a low frequency. Familiarity will be higher*
> *for high frequency tasks than for low frequency tasks.*

Two-way within-subjects ANOVAs were run on the EP-15 tasks from both Session 1 and Session 2 to examine the effects of Session and Prior Frequency on the dependent variables. Table 13 shows the ANOVA design.

Table 13. Within-Subjects ANOVA for Session x Prior Frequency (EP-15 S12 tasks)

Source	*Numerator df*	*Denominator df*
Session (D)	1	16
Prior Freqeuncy (Q)	2	31*
DxQ	2	29*

    * as reported by SAS proc mixed

The results of the ANOVAs are shown in Table 14. The interaction effect of Session x Prior Frequency was significant at $p <= 0.05$ for task completion time, task familiarity, and before-task difficulty. The main effect of Prior Frequency was significant for all measured and dependent variables at $p <= 0.01$. The main effect of Session is consistent with the main effect of Session observed in the Controlled Factors Analysis in Section 5.2.

Table 14. Results of Two-Way ANOVA for Session x Prior Frequency (EP-15 S12 tasks)

	Session (D)		*Prior Freq (Q)*		*Interaction (DxQ)*	
	F	*P*	*F*	*P*	*F*	*P*
Task Completion Time	12.27	<.01*	20.16	<.01*	3.60	0.04**
Number of URLs	7.11	0.02**	9.26	<.01*	2.16	0.13
Number of Searches	3.40	0.08	9.84	<.01*	0.36	0.70
Search Ratio	0.18	0.68	9.43	<.01*	0.39	0.68
Task Familiarity	43.76	<.01*	43.07	<.01*	8.29	<.01*
Before-task Difficulty	35.67	<.01*	45.46	<.01*	3.30	0.05**
After-task Difficulty	22.32	<.01*	20.48	<.01*	2.71	0.08

$* p \leq 0.01, ** p \leq 0.05$

5.3.1   Effects of Session x Prior Frequency Interaction

The Session x Prior Frequency interaction had a significant effect at $p <= 0.01$ on task familiarity and at $p <= 0.05$ on task completion time and before-task difficulty. Post-hoc analyses using the Tukey-Kramer adjustment were conducted to examine the Session x Prior Frequency interaction effect on task completion time, task familiarity, and before-task difficulty. Note that in the graphs, since *prior* frequency (i.e. frequency as measured in the before-task questions of session 1) is being used, the columns consider the same tasks across the two days.
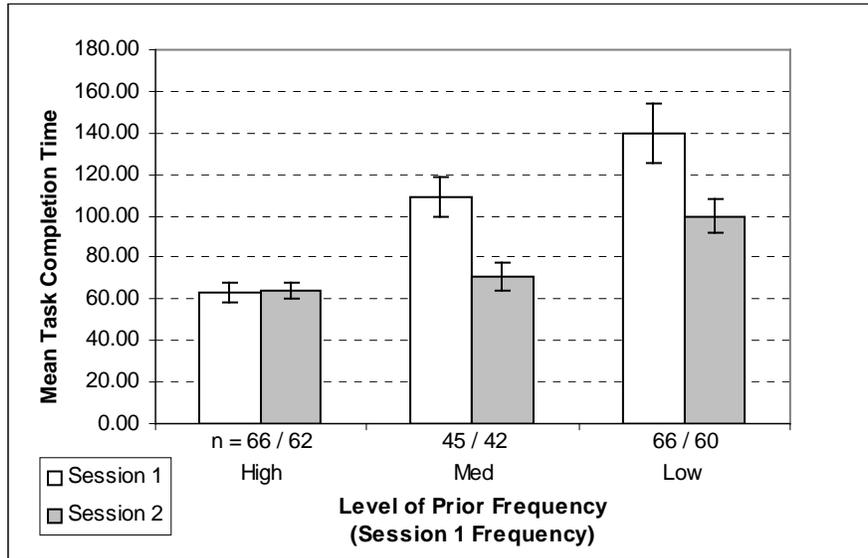
Figure 28. Mean Task Completion Time for Session x Prior Frequency (EP-15 tasks)

Table 15. Post-hoc comparisons of Mean Task Completion Time for Session x Prior Frequency (EP-15 tasks)

		Session 1			*Session 2*		
		High	*Med*	*Low*	*High*	*Med*	*Low*
Session 1	*High*		0.02	<.01	1.00	0.99	0.07
	Med			0.26	0.02	0.11	0.96
	Low				<.01	<.01	0.03
Session 2	*High*					0.99	0.08
	Med						0.35
	Low						
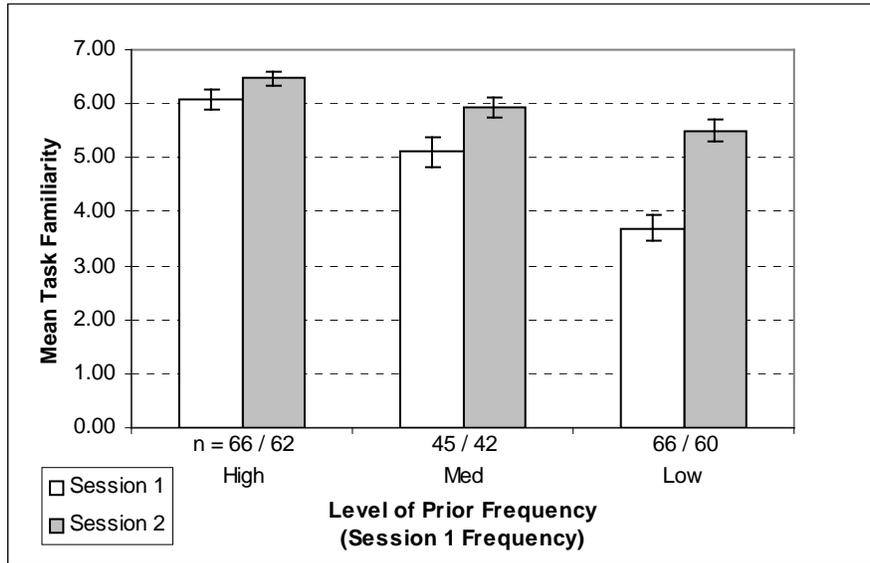
Figure 29. Mean Task Familiarity for Session x Prior Frequency (EP-15 tasks)

Table 16. Post-hoc comparisons of Mean Task Familiarity for Session x Prior Frequency (EP-15 tasks)

		Session 1			*Session 2*		
		High	*Med*	*Low*	*High*	*Med*	*Low*
Session 1	*High*		0.01	<.01	0.60	0.98	0.16
	Med			<.01	<.01	0.08	0.71
	Low				<.01	<.01	<.01
Session 2	*High*					0.34	0.01
	Med						0.61
	Low						
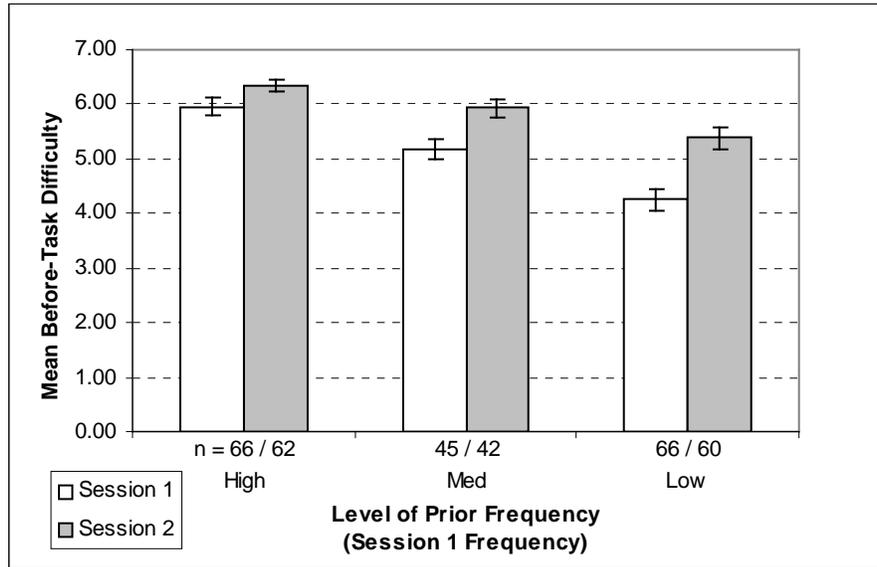
Figure 30. Mean Before-Task Difficulty for Session x Prior Frequency (EP-15 tasks)

Table 17. Post-hoc comparisons of Mean Before-Task Difficulty for Session x Prior Frequency (EP-15 tasks)

		Session 1			Session 2		
		High	Med	Low	High	Med	Low
Session 1	High		<.01	<.01	0.46	0.99	0.02
	Med			<.01	<.01	0.04	0.97
	Low				<.01	<.01	<.01
Session 2	High					0.29	<.01
	Med						0.15
	Low						

5.3.2   Effects of Session

Session had a significant main effect at $p<=0.02$ on the dependent variables task completion time, number of URLs, task familiarity, before-task difficulty, and after-task difficulty. These differences are consistent with the analysis of the effects of Session in the controlled factors analysis in Section 5.2.3.

5.3.3   Effects of Prior Frequency

Prior Frequency had a significant main effect at $p<=0.01$ on all the dependent variables: task completion time, number or URLs, number of searches, search ratio, task familiarity, task frequency (see note about task frequency in section 4.4.4.1), before-task difficulty, and after-task difficulty.

Post-hoc analyses using the Tukey-Kramer adjustment were conducted for each of the dependent variables to examine the significant differences of Prior Frequency. Table 18 shows the means and standard error values for each of the dependent variables for each level of Prior Frequency. Figure 31 shows graphs of the mean and standard error values for each of the dependent variables. The post-hoc comparisons are shown in Table 19.

Table 18. Prior Frequency Mean and Standard Error Values

	High		**Med**		**Low**	
	Mean	*SE*	*Mean*	*SE*	*Mean*	*SE*
Task Completion Time	63.55	3.10	90.75	6.38	120.83	8.60
Number of URLs	4.24	0.40	6.34	0.52	7.18	0.57
Number of Searches	0.30	0.07	0.92	0.21	1.10	0.14
Search Ratio	0.06	0.01	0.11	0.02	0.15	0.02
Task Familiarity	6.27	0.11	5.51	0.17	4.56	0.18
Before-task Difficulty	6.14	0.10	5.53	0.13	4.79	0.15
After-task Difficulty	6.36	0.09	5.87	0.14	5.56	0.12
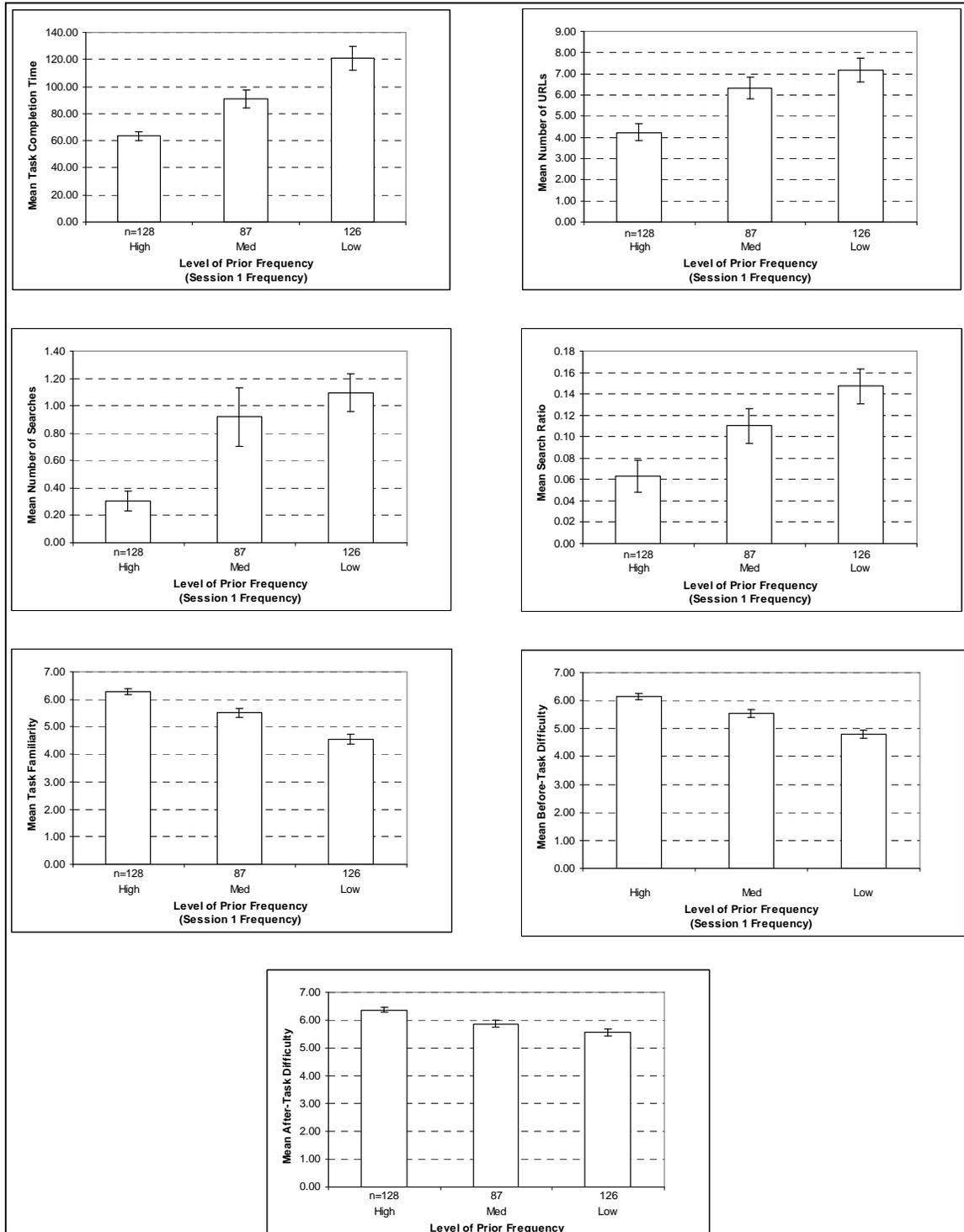
Figure 31. Prior Frequency Effect Mean Values

Table 19. Post-hoc Task Frequency Effects (EP-15 S12 tasks)

	p-value	
Task completion time		
High-Low	<.01	*
High-Medium	0.02	**
Medium-Low	0.02	**
Number of URLs		
High-Low	<.01	*
High-Medium	0.03	**
Medium-Low	0.51	
Number of searches		
High-Low	<.01	*
High-Medium	0.02	**
Medium-Low	0.52	
Search ratio		
High-Low	<.01	*
High-Medium	0.12	
Medium-Low	0.14	
Task familiarity		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	<.01	*
Before-task difficulty		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	<.01	*
After-task difficulty		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	0.05	**

* p ≤ 0.01, ** p ≤ 0.05

5.3.4   Prior Frequency Analysis of Session 1 tasks only

This section continues the analysis of prior frequency by considering the effect of prior frequency on only the EP-15 Session 1 tasks. Session 1 data provides information about participants doing the tasks for the first time as part of this study.

This analysis examines finding and refinding based on the prior frequency with which participants did each task. Only session 1 tasks are included in the analysis. Prior frequency was mapped to three levels: low (less often than once per month), medium (once or several times per month), and high (once per week or more).

One-way within-subjects ANOVAs were run on the EP-15 tasks from Session 1 to examine the effects of prior frequency (H) on the dependent variables. Table 20 shows the ANOVA design.

Table 20. Within-Subjects One-Way ANOVA for Frequency (EP-15 S1 tasks)

Source	Numerator df	Denominator df
Prior Frequency (H)	2	31*

     * as reported by SAS proc mixed

The results of the ANOVAs are shown in Table 21. The main effect of Prior Frequency on the Session 1 EP-15 tasks was significant at $p \leq 0.01$ for: task completion time, number of URLs, task familiarity, before-task difficulty and after-task difficulty. The main effect of Prior Frequency was significant at $p \leq 0.05$ for number of searches and search ratio.

Table 21. Results of One-Way ANOVA for Prior Frequency (EP-15 D1 tasks)

	Num DF	Den DF	Prior Frequency (H) F	P
Task completion time	2	31	15.56	<.01*
Number of URLs	2	31	8.70	<.01*
Number of searches	2	31	4.74	0.02**
Search ratio	2	32	3.21	0.05**
Task familiarity	2	31	33.59	<.01*
Before-task difficulty	2	31	29.96	<.01*
After-task difficulty	2	31	12.34	<.01*

* p ≤ 0.01, ** p ≤ 0.05

Post-hoc analyses using the Tukey-Kramer adjustment were conducted for each of the dependent variables to examine the significant differences of Prior Frequency. Table 22 shows the means and standard error values for each of the dependent variables for each level of Prior Frequency. Figure 32 shows graphs of the mean and standard error values for each of the dependent variables. The post-hoc comparisons are shown in Table 23.
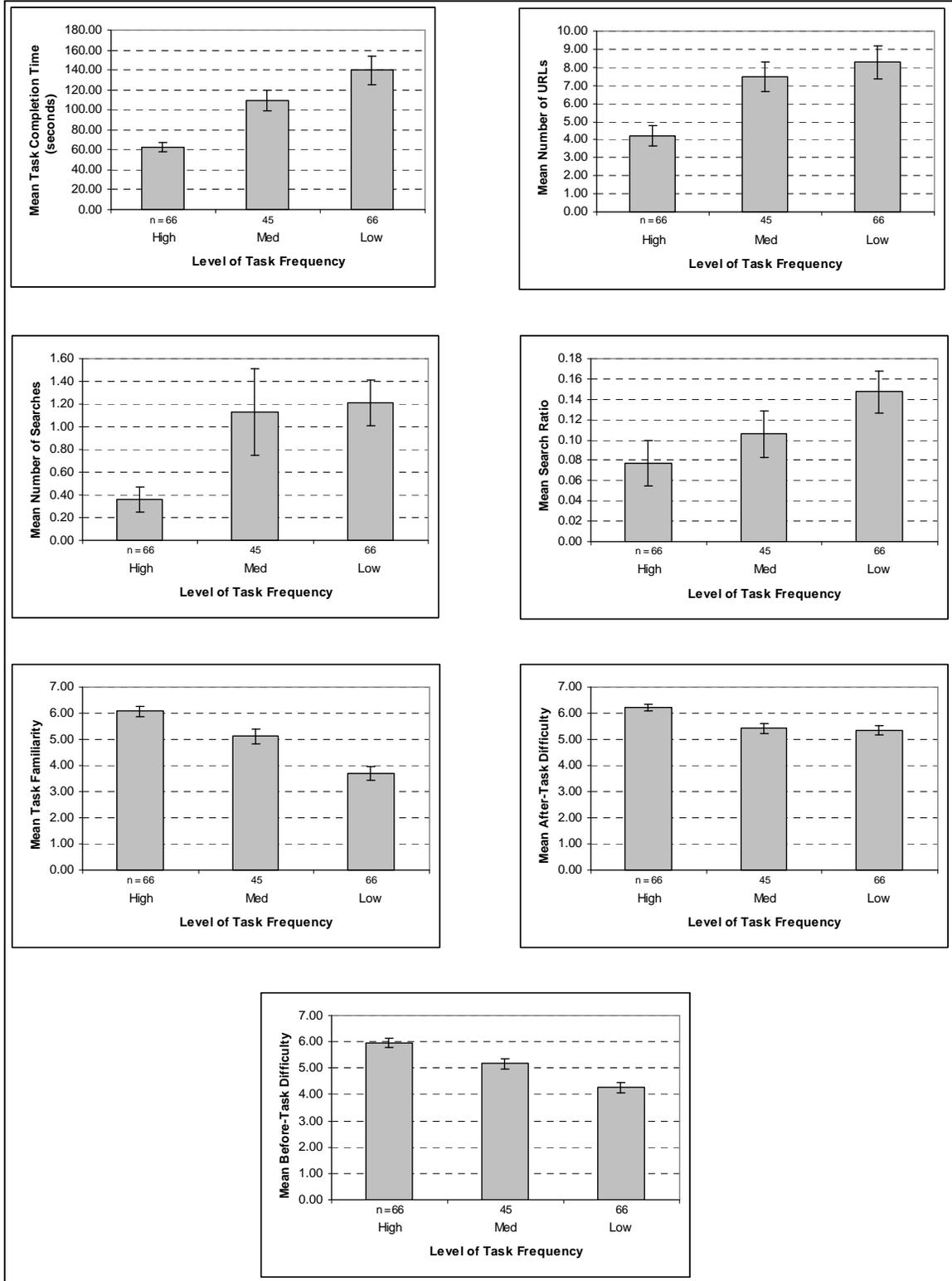
Figure 32. Effect of Prior Frequency on Session 1 only – Means and Standard Error Values (EP-15 S1 tasks)

Table 22. Mean Values for Level of Prior Frequency (EP-15 S1 tasks)

Level of Frequency	High (n=66)		Medium (n=45)		Low (n=66)	
	Mean	SE	Mean	SE	Mean	SE
Task completion time	62.98	4.56	109.20	9.94	139.83	14.23
Number of URLs	4.20	0.57	7.47	0.81	8.29	0.92
Number of searches	0.36	0.11	1.13	0.38	1.21	0.20
Search ratio	0.08	0.02	0.11	0.02	0.15	0.02
Task familiarity	6.08	0.18	5.11	0.27	3.70	0.25
After-task difficulty	6.23	0.14	5.34	0.17	5.42	0.20
Before-task difficulty	5.95	0.16	4.26	0.19	5.18	0.20

Table 23. Post-hoc Task Frequency Effects (EP-15 S1 tasks)

	p-value	
Task completion time		
High-Low	<.01	*
High-Medium	0.02	**
Medium-Low	0.11	
Number of URLs		
High-Low	<.01	*
High-Medium	0.02	**
Medium-Low	0.75	
Number of searches		
High-Low	0.02	**
High-Medium	0.07	
Medium-Low	0.95	
Search ratio		
High-Low	0.04	**
High-Medium	0.70	
Medium-Low	0.32	
Task familiarity		
High-Low	<.01	*
High-Medium	0.01	*
Medium-Low	<.01	*
Before-task difficulty		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	<.01	*
After-task difficulty		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	0.89	

* p ≤ 0.01, ** p ≤ 0.05

The post-hoc comparisons shown in are described in Table 23 and in detail below.

> *Task completion time* – participants completed high-frequency session 1 tasks significantly faster than medium- and low-frequency tasks (p≤0.05), but no significant difference was detected between low and medium frequency tasks.

Number of URLs –high frequency session 1 tasks were completed with significantly fewer URLs than medium and low frequency tasks (p≤0.05), but no difference was detected between low and medium frequency tasks.

Number of searches – high frequency session 1 tasks were completed with fewer searches than low frequency tasks (p≤0.05) or medium frequency tasks (p=0.07). No significant difference was detected between low and medium frequency tasks.

Search Ratio – high frequency session 1 tasks had significantly lower search ratios than low frequency tasks (p≤0.01), but no significant difference was detected between high and medium tasks or between low and medium frequency tasks in terms of search ratio.

Task familiarity – for the session 1 tasks, all three levels of frequency had significantly different values of task familiarity than all the other levels (p≤0.01).

Before-task difficulty – for the session 1 tasks, all three levels of frequency had significantly different values of before-task difficulty than all the other levels (p≤0.01).

After-task difficulty – for the session 1 tasks, high frequency tasks were rated as being less difficult on the after-task difficulty rating than medium and low frequency tasks (p≤0.01), but no significant difference was detected between low and medium frequency tasks.

5.3.5   Summary

In this section, the analysis is summarized in terms of the hypotheses for this analysis.

> *Hypothesis 2 – Finding and Refinding have differences when defined by Task Frequency*
> *There will be a significant effect of task frequency on task completion time, number of URLs, number of searches, search ratio, familiarity, before-task difficulty, and after-task difficulty for the Session 1 and Session 2 tasks. Specifically, task completion time, number of URLs, number of searches, search ratio, before-task difficulty, and after-task difficulty will all be lower/easier for tasks with a high level of frequency than for tasks with a low frequency. Familiarity will be higher for high frequency tasks than for low frequency tasks.*

Session had a significant main effect on task completion time, number of URLs, task familiarity, before-task difficulty, and after-task difficulty at p<=0.02. However, there was also a significant Session x Prior Frequency interaction for task completion time, task familiarity, and before-task difficulty at p<=0.05.

Session x Prior Frequency – Post-hoc tests of the Session x Prior Frequency interaction revealed several interesting results. For session 1, high frequency tasks took significantly less time than medium and low frequency tasks. For session 2, both high and medium frequency tasks took about the same amount of time as high frequency tasks in session 1. Low frequency session 2 tasks took significantly less time than low frequency session 1 tasks, but still took more time than high or medium frequency tasks. These results indicate that one session of experience: did not change the task completion time for high prior frequency tasks, lowered the time of medium frequency tasks to that of high prior frequency tasks, and lowered the time for low prior frequency tasks, but not to the same levels as high and medium prior frequency tasks.

For session 1, high frequency tasks had significantly higher familiarity scores than medium frequency tasks, which also had significantly higher familiarity scores than low frequency tasks. Although the mean familiarity scores for all three levels of prior frequency were higher for session 2 than session 1, only the low prior frequency tasks were significantly higher (p<=0.05). For session 2, the high prior frequency tasks had significantly higher familiarity ratings than the session 2 low prior frequency tasks, but the medium prior frequency tasks were not significantly different from either high or low. These results

indicate that low prior frequency tasks benefited the most from one session of experience, bringing their familiarity rating up to that of the session 1 medium frequency tasks.

Results for the Session x Prior Frequency interaction in terms of the before-task difficulty ratings are similar to the results for the familiarity ratings. For session 1, high, medium, and low prior frequency all had significantly different before-task difficulty ratings – high was rated significantly easier than medium, which was rated significantly easier than low. For session 2, both medium and low prior frequency tasks were rated as significantly easier than they were in session 1. The session 2 low frequency tasks were rated at a level of before-task difficulty that was similar to that of session 1 medium frequency tasks.

Session main effect – For the main effect of Session, there was a significant effect on task completion time, number of URLs, task familiarity, before-task difficulty, and after-task difficulty. The observed effects are similar to those observed in the Session x Task analysis described in Section 5.2.4 on page 70.

Prior frequency main effect – Prior frequency had a significant main effect at $p<=0.01$ on the dependent variables: task completion time, number or URLs, number of searches, search ratio, task familiarity, before-task difficulty, and after-task difficulty. However, as described previously, there was a significant Session x Prior frequency interaction for task completion time, task familiarity, and before-task difficulty at $p<=0.05$.

Post-hoc tests revealed differences based on prior frequency. For task completion time, all three levels of prior frequency had significantly different mean task completion times at $p<=0.05$, with high < medium < low. High prior frequency was significantly different from medium and low in terms of number of URLs and number of searches. For search ratio, only high and low prior frequency were significantly different. For task familiarity, before-task difficulty, and after-task difficulty, all three levels of prior frequency were different from each other at $p<=0.05$.

Conclusion – the hypothesis is mainly supported with some additional observations. Prior frequency did have a significant main effect at $p<=0.01$ on all the dependent variables. For all dependent measures, high prior frequency was significantly different from low prior frequency at $p<=0.05$ and for several of the measures, high, medium, and low were all different from each other with high < medium < low. However, the interaction of Session x Prior Frequency also had a significant effect at $p<=0.05$ on task completion time, task familiarity, and before-task difficulty. These results showed that, in general, the experience of having done the task previously improved the measures for medium and low prior frequency tasks more than for high prior frequency tasks.

5.4   Analysis 3 – Prior Familiarity Analysis

Prior familiarity is the familiarity for each task as measured *before* the participant performed the task during Session 1. This analysis examines the effects of factors session (D) and prior familiarity (K) on the dependent variables. Data from both sessions is included in this analysis, but prior familiarity is the task familiarity as measured during the first session, before the participant had performed the task. Because this analysis compares data across sessions, only the EP-15 tasks are included (these are the exact (E) and path (P) tasks except for task 15).

This analysis addresses Hypothesis 4:

Hypothesis 4 – Task familiarity affects finding and refinding
There will be a significant effect of prior task familiarity on the dependent measures task
completion time, number of URLs, number of searches, search ratio, frequency, before-task
difficulty, and after-task difficulty for the Session 1 and Session 2 tasks. Specifically, task
completion time, number of URLs, number of searches, search ratio, before-task difficulty, and
after-task difficulty will all be lower/easier for tasks with a high level of prior familiarity than
tasks with low familiarity. Frequency will be higher for high prior familiarity tasks than for low
prior familiarity tasks.

Two-way within-subjects ANOVAs were run on the EP-15 tasks from both Session 1 and Session 2 to
examine the effects of Session and Prior Familiarity on the dependent variables. Table 24 shows the
ANOVA design.

Table 24. Within-Subjects ANOVA for Session x Prior Familiarity (EP-15 S12 tasks)

Source	*Numerator df*	*Denominator df*
Session (D)	1	16
Prior Familiarity (K)	2	27*
DxK	2	27*

    *as reported by SAS proc mixed

The results of the ANOVAs are shown in Table 14. The interaction effect of Session x Prior Familiarity
was significant at p<= 0.05 for task completion time, number of URLs, before-task difficulty, and after-task
difficulty. The main effect of Prior Familiarity was significant for all measured and dependent variables.
The main effect of Session was mainly consistent with the main effect of Session observed in the
Controlled Factors Analysis in Section 5.2.

Table 25. Results of Two-Way ANOVA for Session x Prior Familiarity (EP-15 S12 tasks)

	Session (D)		*Prior Fam (K)*		*Interaction (DxK)*	
	F	*P*	*F*	*P*	*F*	*P*
Task Completion Time	13.29	<.01*	9.06	<.01*	4.25	0.02**
Number of URLs	7.56	0.01*	4.81	0.02**	5.18	0.01*
Number of Searches	4.80	0.04**	24.43	<.01*	1.71	0.20
Search Ratio	1.13	0.30	50.18	<.01*	1.28	0.29
Task Frequency~	0.59	0.45	65.37	<.01*	0.97	0.39
Before-task Difficulty	81.14	<.01*	141.39	<.01*	23.21	<.01*
After-task Difficulty	20.50	<.01*	26.58	<.01*	9.09	<.01*

* $p \le 0.01$, ** $p \le 0.05$, ~ see note about task frequency in section 4.4.4.1

5.4.1   Effects of Session x Prior Familiarity Interaction

The Session x Prior Familiarity interaction had a significant effect at p<=0.02 on task completion time,
number of URLs, before-task difficulty and after-task difficulty. Post-hoc analyses using the Tukey-
Kramer adjustment were conducted to examine the Session x Prior Familiarity interaction effect on task
completion time, number of URLs, before-task difficulty and after-task difficulty. Note that in the graphs,
since *prior* familiarity (i.e. familiarity as measured in the before-task questions of session 1) is being used,
the columns consider the same tasks across the two days.
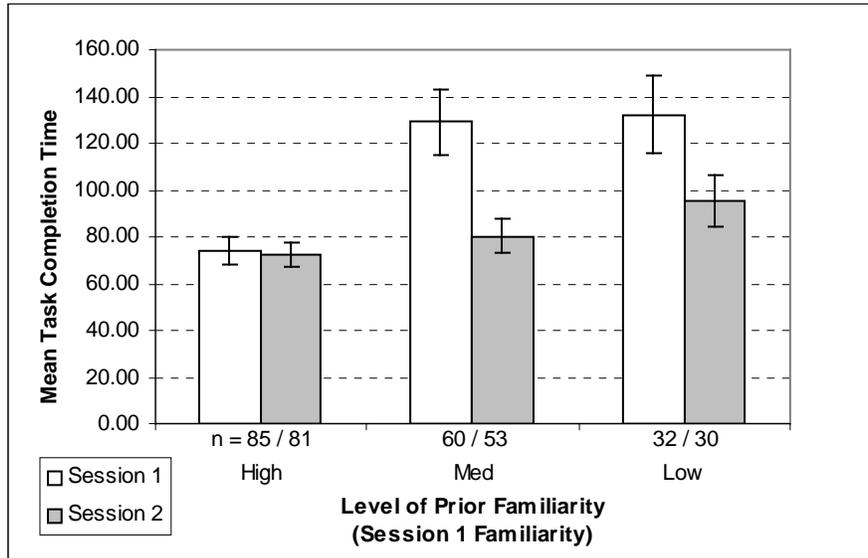
Figure 33. Mean Task Completion Time for Session x Prior Familiarity (EP-15 S12 tasks)

Table 26. Post-hoc comparisons of Mean Task Completion Time for Session x Prior Familiarity (EP-15 S12 tasks)

		Session 1			*Session 2*		
		High	*Med*	*Low*	*High*	*Med*	*Low*
Session 1	*High*		<.01	0.01	1.0	1.0	0.83
	Med			1.0	<.01	0.01	0.27
	Low				0.01	0.03	0.31
Session 2	*High*					1.0	0.76
	Med						0.94
	Low						
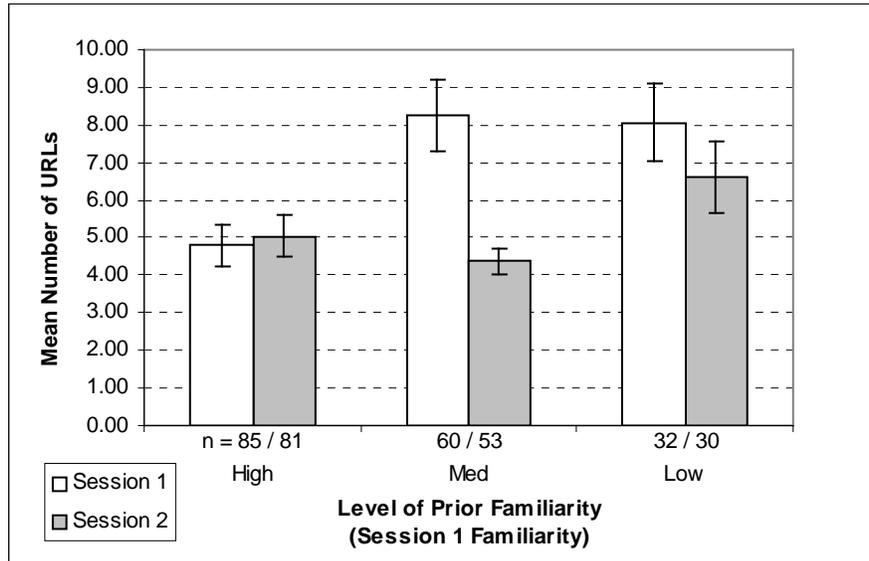
Figure 34. Mean Number of URLs for Session x Prior Familiarity (EP-15 S12 tasks)

Table 27. Post-hoc comparisons of Mean Number of URLs for Session x Prior Familiarity (EP-15 S12 tasks)

		Session 1			*Session 2*		
		High	*Med*	*Low*	*High*	*Med*	*Low*
Session 1	*High*		0.01	0.07	1.0	0.99	0.64
	Med			1.0	0.02	0.01	0.76
	Low				0.12	0.04	0.89
Session 2	*High*					0.96	0.77
	Med						0.44
	Low						
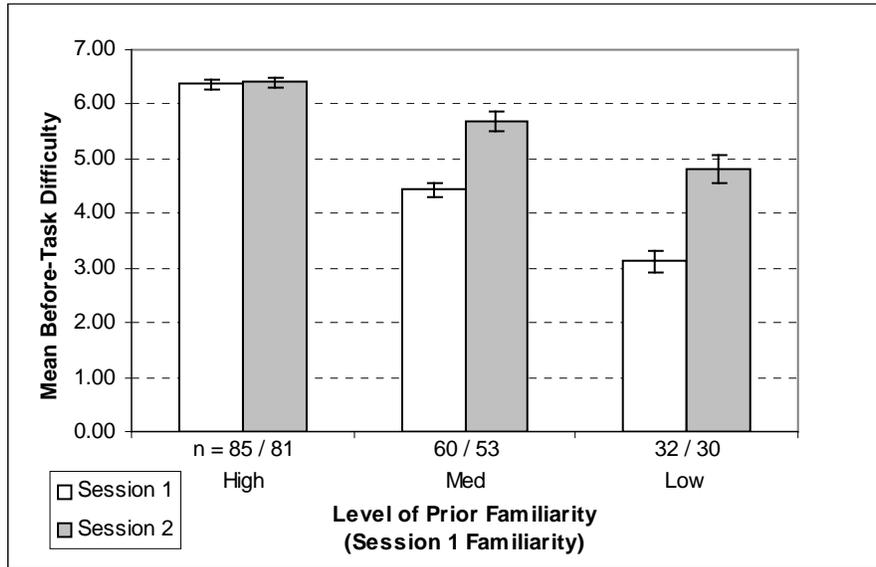
Figure 35. Mean Before-Task Difficulty for Session x Prior Familiarity (EP-15 S12 tasks)

Table 28. Post-hoc comparisons of Mean Before-Task Difficulty for Session x Prior Familiarity (EP-15 S12 tasks)

		Session 1			Session 2		
		High	*Med*	*Low*	*High*	*Med*	*Low*
Session 1	*High*		<.01	<.01	1.0	0.04	<.01
	Med			<.01	<.01	<.01	0.80
	Low				<.01	<.01	<.01
Session 2	*High*					0.02	<.01
	Med						<.01
	Low						
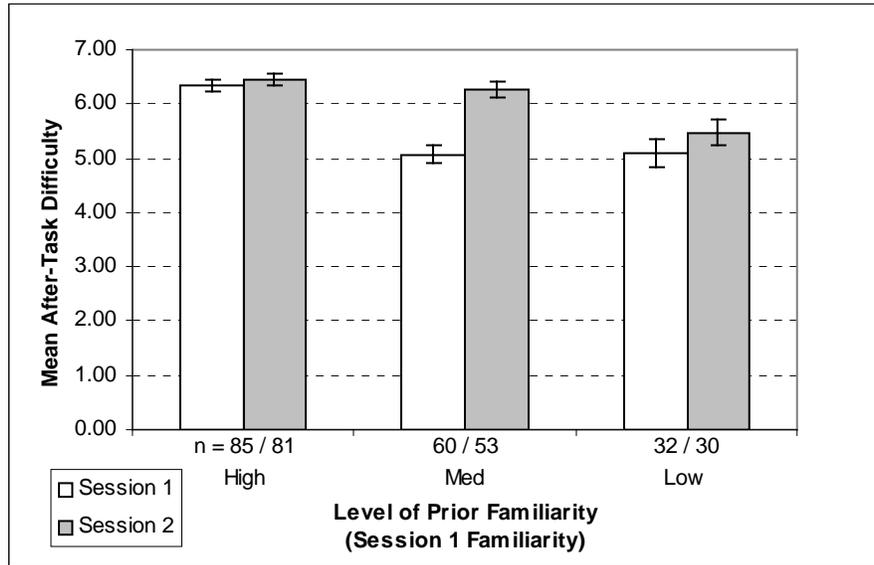
Figure 36. Mean After-Task Difficulty for Session x Prior Familiarity (EP-15 S12 tasks)

Table 29. Post-hoc comparisons of Mean After-Task Difficulty for Session x Prior Familiarity (EP-15 S12 tasks)

		Session 1			Session 2		
		High	Med	Low	High	Med	Low
Session 1	High		<.01	<.01	0.99	1.0	0.01
	Med			1.0	<.01	<.01	0.82
	Low				<.01	<.01	0.70
Session 2	High					1.0	<.01
	Med						0.02
	Low						

5.4.2   Effects of Session

Session had a significant main effect at $p<=0.01$ on the dependent variables task completion time, number of URLs, before-task difficulty, and after-task difficulty. At $p<=0.05$, it had a significant main effect on number of searches. These differences are similar to those found in Section 5.2.3.

5.4.3   Effects of Prior Familiarity

Prior Familiarity had a significant main effect at $p<=0.02$ on the dependent variables: task completion time, number or URLs, number of searches, search ratio, task frequency (see note about task frequency in section 4.4.4.1), before-task difficulty, and after-task difficulty.

Post-hoc analyses using the Tukey-Kramer adjustment were conducted for each of the dependent variables to examine the significant differences of Prior Familiarity. Table 30 shows the means and standard error values for each of the dependent variables for each level of Prior Familiarity. Figure 37 shows graphs of the mean and standard error values for each of the dependent variables. The post-hoc comparisons are shown in Table 31.

Table 30. Prior Familiarity Mean and Standard Error Values (EP-15 S12 tasks)

	High, n=166		**Med**, n=113		**Low**, n=62	
	Mean	*SE*	*Mean*	*SE*	*Mean*	*SE*
Task Completion Time	73.25	3.88	106.21	8.46	114.40	10.35
Number of URLs	4.92	0.39	6.42	0.56	7.37	0.70
Number of Searches	0.29	0.07	0.93	0.17	1.68	0.22
Search Ratio	0.04	0.01	0.13	0.02	0.25	0.03
Task Frequency~	4.33	0.13	5.94	0.13	6.55	0.20
Before-task Difficulty	6.39	0.07	5.03	0.12	3.94	0.19
After-task Difficulty	6.40	0.08	5.63	0.12	5.27	0.17

~ see note about task frequency in section 4.4.4.1
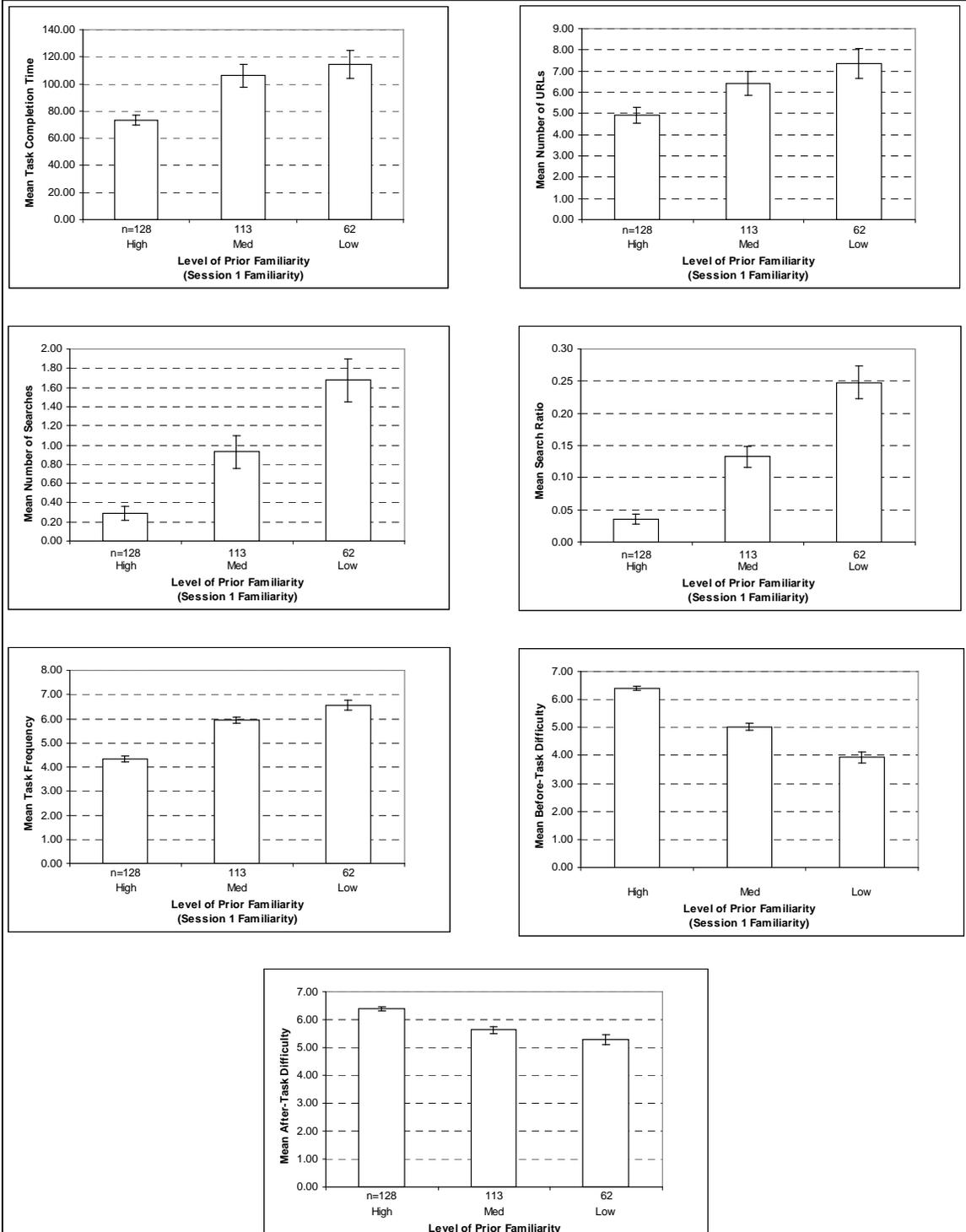
**Figure 37. Prior Familiarity Effects Mean Values (EP-15 S12 tasks)**

Table 31. Post-hoc Task Familiarity Effects (EP-15 S12 tasks)

	p-value	
Task completion time		
High-Low	<.01	*
High-Medium	0.01	*
Medium-Low	0.74	
Number of URLs		
High-Low	0.02	**
High-Medium	0.15	
Medium-Low	0.44	
Number of searches		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	<.01	*
Search ratio		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	<.01	*
Task frequency		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	<.01	*
Before-task difficulty		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	<.01	*
After-task difficulty		
High-Low	<.01	*
High-Medium	<.01	*
Medium-Low	0.03	**

* p ≤ 0.01, ** p ≤ 0.05

5.4.4   Summary

In this section, the analysis is summarized in terms of the hypotheses for this analysis.

> *Hypothesis 4 – Task familiarity affects finding and refinding*
> *There will be a significant effect of prior task familiarity on the dependent measures task completion time, number of URLs, number of searches, search ratio, frequency, before-task difficulty, and after-task difficulty for the Session 1 and Session 2 tasks. Specifically, task completion time, number of URLs, number of searches, search ratio, before-task difficulty, and after-task difficulty will all be lower/easier for tasks with a high level of prior familiarity than tasks with low familiarity. Frequency will be higher for high prior familiarity tasks than for low prior familiarity tasks.*

Session had a significant main effect on task completion time, number of URLs, number of searches, before-task difficulty, and after-task difficulty at p<=0.05. However, there was also a significant Session x Prior Familiarity interaction for task completion time, number of URLs, before-task difficulty, and after-task difficulty at p<=0.02.

Session x Prior Familiarity – Post-hoc tests of the Session x Prior Familiarity interaction revealed several interesting results. For session 1, high prior familiarity tasks took significantly less time than medium and low prior familiarity tasks. For session 2, both high and medium familiarity tasks took about the same amount of time as high familiarity tasks in session 1. Low familiarity session 2 tasks took less time than low frequency session 1 tasks, but not significantly. In addition, none of the three levels of prior familiarity were significantly different for session 2 in terms of mean task completion time. These results indicate that: one session of experience did not change the task completion time for high prior familiarity tasks, significantly lowered the time of medium familiarity tasks to that of high prior frequency tasks, and lowered the time for low prior familiarity tasks, but not significantly.

In terms of the number of URLs, for session 1, high familiarity tasks involved significantly lower numbers of URLs than medium familiarity tasks. For session 2, the number of URLs was not significantly changed for high familiarity tasks from session 1, but was significantly lower for medium familiarity tasks than in session 1, bringing them to about the same level as session 1 high familiarity tasks. Session 2 low familiarity tasks did involve fewer URLs than session 1 low familiarity tasks, but not significantly. There were no significant differences between the three levels for session 2 in terms of number of URLs.

Results for the Session x Prior Familiarity interaction in terms of the before-task difficulty ratings show that for session 1, high, medium, and low prior familiarity all had significantly different before-task difficulty ratings – high was rated significantly easier than medium which was rated significantly easier than low. A similar pattern was observed for session 2 – high < medium < low – with all three being significantly different. The tasks with high prior familiarity did not have a significantly different mean before-task difficulty rating from session 1 to session 2. However, both the medium and low tasks did have significantly different mean before-task difficulty ratings from session 1 to session 2. These results indicate that medium and low prior familiarity tasks benefited from one session of experience in terms of their before-task difficulty ratings.

The after-task difficulty ratings for prior familiarity had slightly different results than the before-task difficulty ratings. Session 1 tasks with high prior familiarity were rated as significantly easier than tasks with medium or low prior familiarity. In session 2, tasks with low prior familiarity were rated as more difficult than tasks with high or medium prior familiarity. However, from session 1 to session 2, only the medium prior familiarity tasks were found to be rated significantly easier in session 2. No difference was observed from session 1 to session 2 for the high and low prior familiarity tasks in terms of after-task difficulty.

Session main effect – For the main effect of Session, there was a significant effect on task completion time, number of URLs, before-task difficulty, and after-task difficulty at $p<=0.01$. The observed effects are similar to those observed in the Session main effect summary described in Section 5.2.4 on page 70, except that in this analysis, at $p<=0.05$, there was a main effect of Session on the number of searches.

Prior familiarity main effect – Prior familiarity had a significant main effect at $p<=0.02$ on the dependent variables: task completion time, number of URLs, number of searches, search ratio, before-task difficulty, and after-task difficulty. However, as described previously, there was a significant Session x Prior Familiarity interaction for task completion time, number of URLs, before-task difficulty, and after-task difficulty at $p<=0.02$.

Post-hoc tests revealed differences based on prior familiarity. For task completion time, high prior familiarity tasks had significantly lower mean task completion times than medium and low prior familiarity tasks at $p<=0.01$. High prior familiarity was significantly different at $p<=0.02$ from low in terms of number of URLs. For number of searches, search ratio, before-task difficulty, and after-task difficulty, all three levels of prior frequency were different from each other at $p<=0.01$ (except for the after-task difficulty comparison of medium to low, which was significant at $p<=0.05$).

5.5   Analysis 4 – Task Similarity Analysis

Task similarity refers to the similarity of the task from Session 1 to Session 2 and is explained in detail in section 4.4.2.5 on page 43. Two analyses of task similarity are presented in this section. First, exact and path tasks (EP-15 tasks) are compared using two-way ANOVAs with Session and Task Similarity as the independent variables. The analysis of the EP-15 S12 tasks is presented in section 5.5.1.

Second, the exact, path, and subset tasks (EPS-15 tasks) are compared for session 2 only using one-way ANOVAs with Task Similarity as the independent variable. Since the subset tasks requested participants to find two items in session 1, but one item in session 2, it is difficult to fairly compare them across session on measures such as task completion time, and number of URLs. However, in session 2, the subset tasks requested that participants refind one of the two items found in session 1 and can be more fairly compared to the exact and path refinding tasks, which also typically asked participants to refind a single item. As such, this analysis using only the session 2 data provides a way to see if there is an effect of task similarity on refinding. The moved task is not included in this analysis because of the low rate of successful completion. The analysis of the EPS-15 tasks for session 2 is presented in section 5.5.2.

These analyses both address Hypotheses 5:

> *Hypothesis 5 – Task similarity affects refinding*
> *There will be a significant effect of task similarity on the dependent measures task completion time, number of URLs, number of searches, search ratio, frequency, before-task difficulty, and after-task difficulty on the Session 2 (refinding) tasks. Specifically, task completion time, number of URLs, number of searches, search ratio, before-task difficulty, and after-task difficulty will all be higher for moved and subset tasks than for exact or path tasks. No difference will be observed for any of the dependent measures between the exact and path tasks. In other words: Refinding will be easier if the relationship of the finding and refinding tasks are "exactly the same" or "on the same path". Refinding will be more difficult if the relationship is "subset" or "moved".*

5.5.1   Task Similarity Analysis of Exact and Path Task for Sessions 1 and 2

Two-way within-subjects ANOVAs were run on the EP-15 tasks from both Session 1 and Session 2 to examine the effects of Session and Task Similarity on the dependent variables. Table 32 shows the ANOVA design.

This analysis examines the effects of factors Session (D) and Task similarity (I) on the dependent variables. Data from both sessions is included in this analysis, but each task has a fixed value of task similarity that applies for both sessions. Because this analysis compares data across sessions, only the EP-15 tasks are included (these are the exact (E) and path (P) tasks except for task 15).

Table 32. Within-Subjects ANOVA for Session x Task Similarity (EP-15 S12 tasks)

Source	*Numerator df*	*Denominator df*
Session (D)	1	16
Task Similarity (I)	1	16*
DxI	1	16*

*as reported by SAS proc mixed

The results of the ANOVAs are shown in Table 33. No significant interaction effect of Session x Task Similarity was observed. The main effect of Task Similarity was significant at $p<=0.01$ for task completion

time, number of searches, search ratio, and task familiarity and at p<=0.05 for before-task difficulty. The main effect of Session was mainly consistent with the main effect of Session observed in the Controlled Factors Analysis in Section 5.2.

Table 33. Results of Two-Way ANOVA for Session x Task Similarity (EP-15 S12 tasks)

	Session (D)		Similarity (I)		Interaction (DxI)	
	F	P	F	P	F	P
Task Completion Time	12.27	<.01*	8.08	0.01*	1.64	0.22
Number of URLs	7.15	0.02**	0.40	0.54	0.89	0.36
Number of Searches	2.24	0.15	11.68	<.01*	0.27	0.61
Search Ratio	0.09	0.76	35.94	<.01*	0.46	0.51
Task Familiarity	28.21	<.01*	7.80	0.01*	2.64	0.12
Task Frequency~	0.04	0.84	0.79	0.39	0.40	0.54
Before-task Difficulty	22.82	<.01*	5.13	0.04**	1.25	0.28
After-task Difficulty	17.05	<.01*	1.34	0.26	0.29	0.59

* p ≤ 0.01, ** p ≤ 0.05, ~ see note about task frequency in section 4.4.4.1

5.5.1.1   *Effects of Session x Task Similarity Interaction*

No significant Session x Task similarity interaction effects were observed.

5.5.1.2   *Effects of Session*

Session had a significant main effect at p<=0.02 on the dependent variables task completion time, number of URLs, task familiarity, before-task difficulty, and after-task difficulty. These differences are similar to those found in for the effects of Session found in Section 5.2.3 and are examined in detail on page 69.

5.5.1.3   *Effects of Task Similarity*

Task similarity had a significant main effect at p<=0.01 on task completion time, number of searches, search ratio, task familiarity, and at p<=0.05 on before-task difficulty.

Post-hoc analyses using the Tukey-Kramer adjustment were conducted for each of the dependent variables to examine the significant differences of Task Similarity. Table 34 shows the means and standard error values for each of the dependent variables for each level of Prior Familiarity, along with the p-values for the post-hoc comparisons. Figure 38 shows graphs of the mean and standard error values for each of the dependent variables.
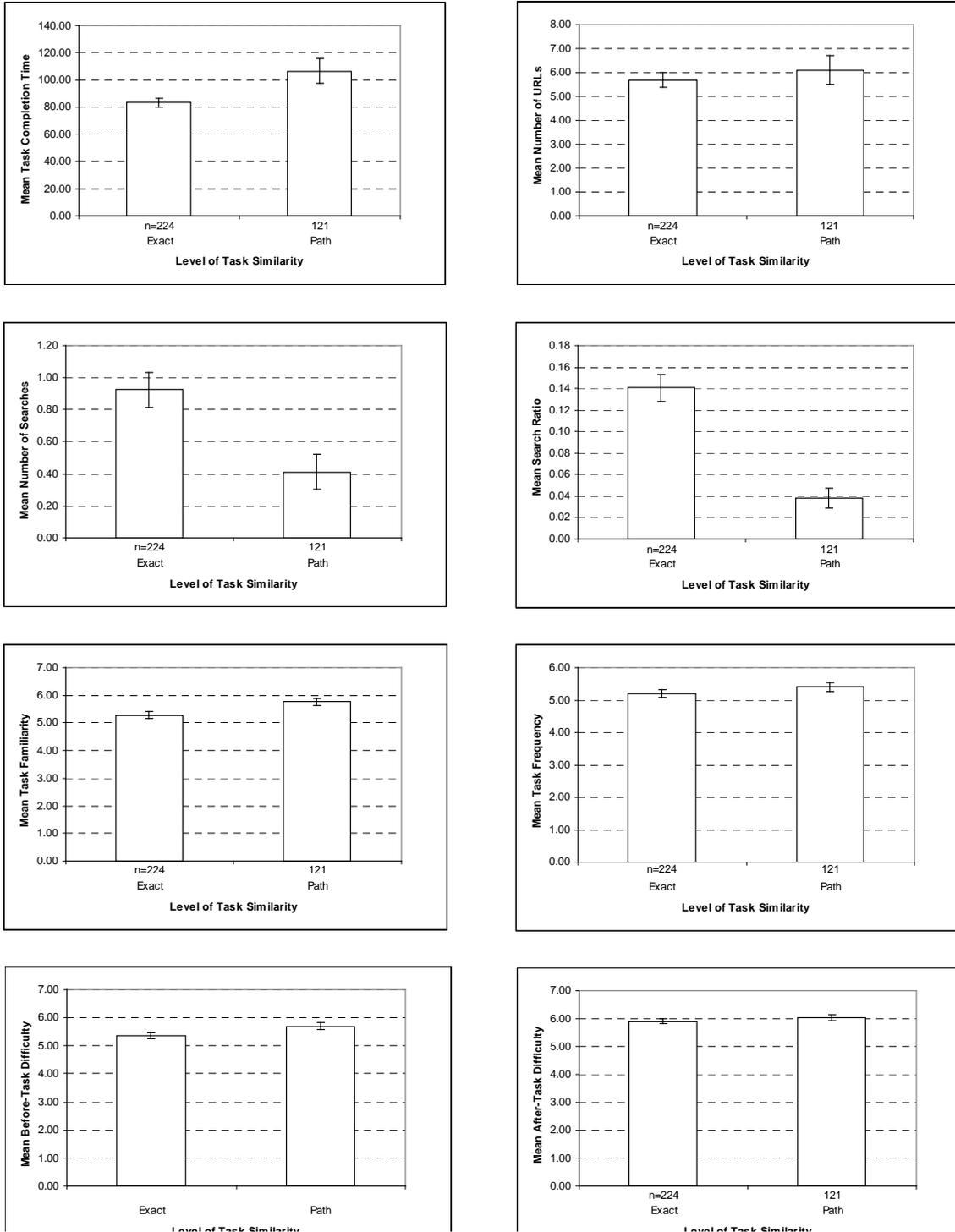
Figure 38. Task Similarity Effect Mean Values (EP-15 S12 tasks)

Table 34. Task Similarity Mean and Standard Error Values (EP-15 S12 tasks)

	Exact, n=224		**Path**, n=121		
	Mean	*SE*	*Mean*	*SE*	*p-val*
Task Completion Time	83.34	3.52	106.42	9.03 *	0.01
Number of URLs	5.69	0.31	6.10	0.62	0.54
Number of Searches	0.92	0.11	0.41	0.11 *	<.01
Search Ratio	0.14	0.01	0.04	0.01 *	<.01
Task Familiarity	5.28	0.13	5.76	0.14 *	0.01
Task Frequency	5.19	0.13	5.40	0.14	0.39
Before-task Difficulty	5.38	0.10	5.69	0.12 **	0.04
After-task Difficulty	5.90	0.09	6.03	0.11	0.26

5.5.2   Task Similarity Analysis of Session 2 tasks only

This analysis includes the exact, path, and subset tasks (EPS-15 tasks) for session 2 only. One-way within-subjects ANOVAs were run on the EPS-15 tasks from Session 2 to examine the effects of Task Similarity on the dependent variables. Table 35 shows the ANOVA design.

Table 35. Within-Subjects One-Way ANOVA for Task Similarity (EPS-15 S2 tasks)

Source	*Numerator df*	*Denominator df*
Task similarity (I)	2	32

The results of the ANOVAs are shown in Table 36. The main effect of Task Similarity on the Session 2 EPS-15 tasks was significant at p≤0.01 for: search ratio, task familiarity, before-task difficulty and after-task difficulty and at p<=0.05 for task completion time and number of searches.

Table 36. Results of One-Way ANOVA for Task Similarity (EPS-15 S2 tasks)

	Num DF	*Den DF*	*Similarity (I)*	
			F	*P*
Task completion time	2	32	4.15	0.03**
Number of URLs	2	32	0.61	0.55
Number of searches	2	32	3.67	0.04**
Search ratio	2	32	6.97	<.01*
Task familiarity	2	32	12.31	<.01*
Before-task difficulty	2	32	10.50	<.01*
After-task difficulty	2	32	9.11	<.01*

*$p \leq 0.01$, ** $p \leq 0.05$*

Post-hoc analyses using the Tukey-Kramer adjustment were conducted for each of the dependent variables to examine the significant differences of Task Similarity. Figure 39 shows graphs of the mean and standard error values for each of the dependent variables. Table 37 shows the means and standard error values for each of the dependent variables for each level of Task Similarity. The post-hoc comparisons are shown in Table 38.
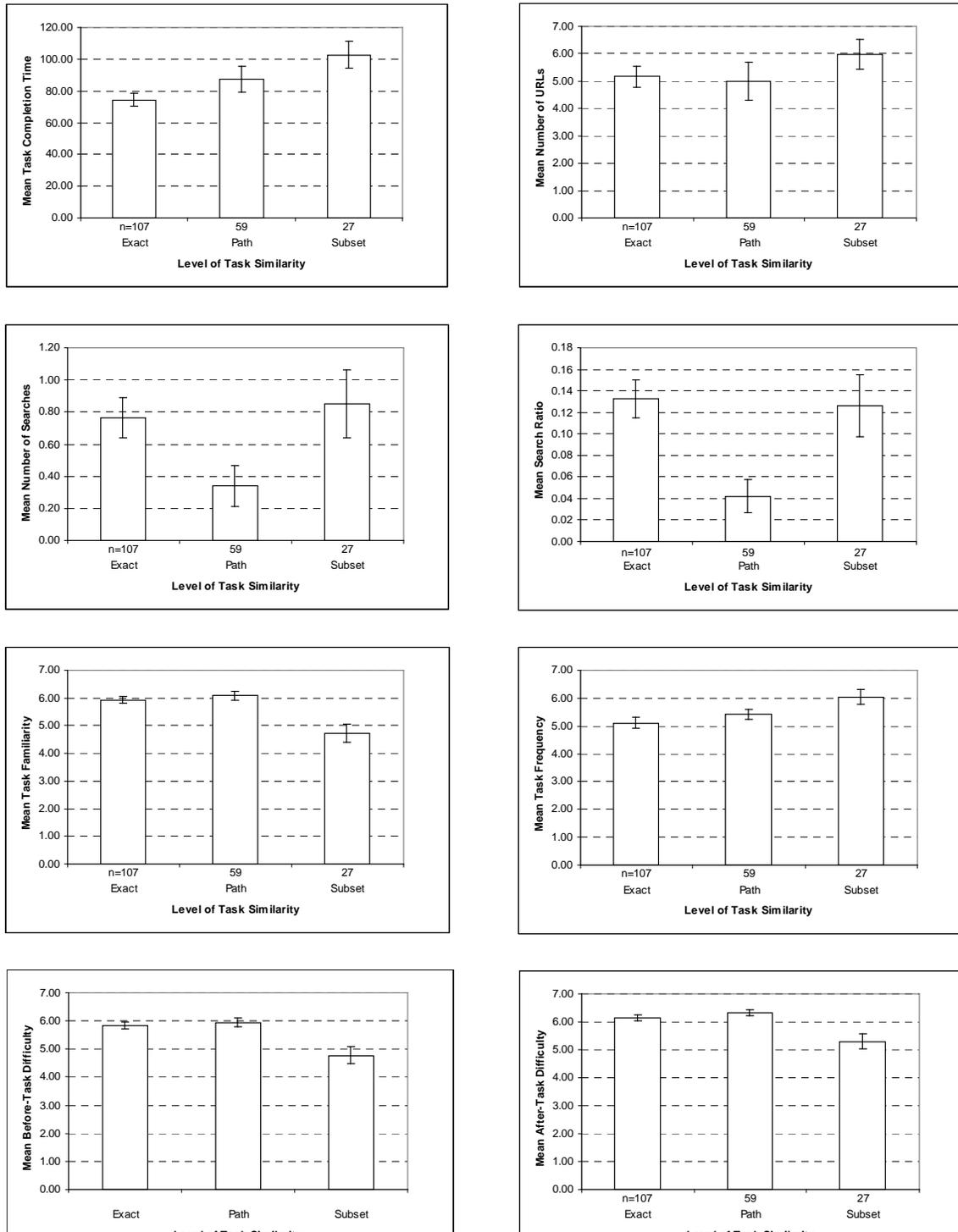
Figure 39. Effect of Task Similarity – Means and Standard Error Values (EPS-15 S2 tasks)

Table 37. Mean Values for Level of Task Similarity (EPS-15 S2 tasks)

	Exact (n=107)		Path (n=59)		Subset (n=27)	
	Mean	*SE*	*Mean*	*SE*	*Mean*	*SE*
Task completion time	74.37	4.20	87.08	8.23	102.59	8.55
Number of URLs	5.16	0.37	4.98	0.69	5.96	0.55
Number of searches	0.77	0.12	0.34	0.12	0.85	0.21
Search ratio	0.13	0.02	0.04	0.02	0.13	0.03
Task familiarity	5.93	0.13	6.08	0.16	4.74	0.33
Task frequency~	5.11	0.18	5.42	0.18	6.04	0.27
After-task difficulty	6.15	0.11	6.33	0.12	5.30	0.28
Before-task difficulty	5.84	0.12	5.95	0.16	4.78	0.31

~ see note about task frequency in section 4.4.4.1

Table 38. Post-hoc Task Frequency Effects (EPS-15 S2 tasks)

	p-value	
Task completion time		
Exact-Path	0.27	
Exact-Subset	0.03	**
Path-Subset	0.33	
Number of URLs		
Exact-Path	0.96	
Exact-Subset	0.61	
Path-Subset	0.54	
Number of searches		
Exact-Path	0.05	**
Exact-Subset	0.88	
Path-Subset	0.10	
Search ratio		
Exact-Path	<.01	*
Exact-Subset	0.99	
Path-Subset	0.05	**
Task frequency		
Exact-Path	0.46	
Exact-Subset	0.05	**
Path-Subset	0.33	
Task familiarity		
Exact-Path	0.53	
Exact-Subset	<.01	*
Path-Subset	<.01	*
Before-task difficulty		
Exact-Path	0.71	
Exact-Subset	<.01	*
Path-Subset	<.01	*
After-task difficulty		
Exact-Path	0.51	
Exact-Subset	<.01	*
Path-Subset	<.01	*

* p ≤ 0.01, ** p ≤ 0.05

5.5.3   Summary

In this section, the analysis is summarized in terms of the hypotheses for this analysis.

> *Hypothesis 5 – Task similarity affects refinding*
> *There will be a significant effect of task similarity on the dependent measures task completion time, number of URLs, number of searches, search ratio, frequency, before-task difficulty, and after-task difficulty on the Session 2 (refinding) tasks. Specifically, task completion time, number of URLs, number of searches, search ratio, before-task difficulty, and after-task difficulty will all be higher for moved and subset tasks than for exact or path tasks. No difference will be observed for any of the dependent measures between the exact and path tasks. In other words: Refinding will be easier if the relationship of the finding and refinding tasks are "exactly the same" or "on the same path". Refinding will be more difficult if the relationship is "subset" or "moved".*

5.5.3.1   *Exact and Path Tasks Across Sessions*

This section summarizes the analysis of the EP-15 tasks across both sessions. Significant main effects were found for Session and Task Similarity. No significant effects of a Session x Task Similarity interaction were observed.

Session main effect – For the main effect of Session, there was a significant effect on task completion time, number of URLs, task familiarity, before-task difficulty, and after-task difficulty at $p <= 0.02$. The observed effects are similar to those observed in the Session x Task analysis described in Section 5.2.4 on page 70.

Task similarity main effect – Task Similarity had a significant main effect on task completion time, number of searches, search ratio, and task familiarity at $p <= 0.01$, and before-task difficulty at $p <= 0.05$. No significant effects of a Session x Task Similarity interaction were observed.

Post-hoc tests revealed significant differences between the exact and path tasks. Exact tasks had a lower mean task completion time, involved larger numbers of searches, had higher search ratio, a lower mean task familiarity rating, and were rated as being harder in the before-task difficulty rating than the path tasks.

5.5.3.2   *Exact, Path, and Subset Tasks in Session 2 only (refinding)*

This section summarizes the analysis of the EPS-15 tasks in session 2 only. Task Similarity had a significant main effect at $p <= 0.01$ on search ratio, task familiarity, before-task difficulty, and after-task difficulty and at $p <= 0.05$ on task completion time and number of searches.

Post-hoc tests showed that subset tasks were rated as significantly more difficult than exact or path tasks in both before- and after-task difficulty ratings. Subset tasks were also rated as significantly less familiar than exact or path tasks. Exact and path tasks were significantly different in terms of the number of searches and search ratio. Path tasks also had a significantly lower search ratio than subset tasks. Exact tasks had a significantly lower task completion time than subset tasks.

5.6   Analysis 6 – Subset and Moved Tasks Analysis

Because of the characteristics of the tasks, the subset (S) and moved (M) tasks were not included in most of the previous analyses. The subset tasks were included in the "Task Similarity Analysis of Session 2 tasks only" presented in section 5.5.2. This section will describe additional analyses of the subset and moved tasks.

5.6.1   Subset Tasks Analysis

There were two subset tasks – task 4 (shopping for sweatshirts), and task 7 (San Francisco tourist attractions). Each task required participants to find two items of interest in session 1 and then refind one of these items in session 2. Table 39 gives the descriptions for each task.

Table 39. Subset Task Descriptions

	Session 1	Session 2
Task 4	Look for prices on Women's Sweatshirts. Find two possibilities that you might buy for a gift to a friend or relative. Please look at at least three different sweatshirts from three different places to buy.	Find one of the Women's Sweatshirts that you looked at during the session last week.
Task 7	Imagine that you are planning a trip to San Francisco, California. Find two tourist attractions or events that you might be interested in. Please look at at least four different attractions or events.	Find one of the tourist attractions or events for San Francisco, California that you looked at during the session last week.

Task 4 was completed successfully by 16 of the 17 participants in session 1. In session 2, 13 of the 17 participants completed it successfully.

Task 7 was completed successfully by 16 of the 17 participants in session 1. In session 2, 15 of the 17 participants completed it successfully.

However, as discussed in section 5.5.2, post-hoc tests showed that subset tasks were rated as significantly more difficult than exact or path tasks in both before- and after-task difficulty ratings. Subset tasks were also rated as significantly less familiar than exact or path tasks.

In these tasks, participants were allowed to select which item to refind. Based on participant comments, many appear to have selected the item for which they had the strongest memory. In addition, many of the participants had considerable domain knowledge of shopping on-line for clothing. Many participants knew the URLs of Web sites for particular brands of clothing and commented that they frequently made purchases from these sites and/or knew them to be of good quality. These factors may have made refinding easier than it would be in different circumstances.

5.6.2   Moved Task Analysis

There was only one moved task in this study – task 2 (news article). This task required participants to find and read a news article of interest and then refind the same news article in session 2. Table 40 gives the descriptions for the task for each session.

Table 40. Moved Task Descriptions

	Session 1	Session 2
Task 2	Find and READ news story about world news	Find the news story that you read last week about world news.

Task 2 was completed successfully by 15 of the 17 participants in session 1. Only 2 of the 17 participants completed it successfully in session 2.

5.7   Analysis 7 – Task-by-Task Analysis

This analysis examines some of the tasks on an individual basis to explore what resources and strategies were commonly used. Analysis of each task is presented in the sections below. Trends and summarizations of these analyses are presented in the discussion in section 6.4.

For each of these task-by-task analyses, only tasks that were attempted, completed successfully, and for which there were not problems with the performance of the task are included in the detailed analysis. Session 2 tasks that were not attempted and completed during the first session are not included in the detailed analysis. This is slightly different than the statistical analyses presented in the previous sections – in those analyses, more stringent criteria were used that included a condition that the before task questions had to have been answered before doing the task.

In the task-by-task analyses presented in this section, tables are presented for each task with information about: the number of participants that attempted the task, the number of participants that successfully completed the task, common resources used, tabulations of the starting strategy used, and an comparison of the starting strategy and resources used in session 1 and session 2. Several of these measures are described in more detail below.

Common Resources Used
This measure lists the top resources that participants appeared to use as the final source for the information and indicates how many participants used each resource.

Starting strategy
This measure indicates the strategy that participants used to start the task and was coded as one of three classifications:

- Go directly – go directly to a known starting point. This does not have to be the final destination of the information, but is an indication that the participant developed some plan other than to use a Web search engine to find the information. The idea of this classification is that the participant appeared to have a starting plan for locating the information that involved going to a known Web site that was not a Web search engine. The participant did not need to know the full path to the information, but did need to exhibit behavior that indicated a starting plan. The most common way this was observed were cases when a participant started a task by typing a URL into the Web browser.

- Search for a resource – use a Web search engine (SE) to try to locate an on-line resource that will contain or lead to the information. The idea here is that the participant did not seem to recall a way to get to the resource, or was unsure of exactly what the name of the resource would be, but used that a search engine to locate a resource. An example is going to a Web search engine and searching for "yellow pages" as a first step to finding the phone number for a local business.

o   Example #1: If the task is to find the Web page of Dr. Stasko at Georgia Tech, then if the participant searched Google for "Georgia Tech", then it would be coded as "SE resource" since the participant was using the SE to locate the main Web page for Georgia Tech (a resource), not the specific information asked for in the task.

o   Example #2: For the task to find the score of the most recent Atlanta Falcons football game, a participant who searched Google for "Espn Sports Score" was classified as searching for a resource. A participant who used an SE for "atlanta falcons score" was classified as searching for the information. A more confusing case was a participant who used an SE to search for "Atlanta Falcons". It was decided to classify this as a search for a resource since the specific information type "score" was not included in the search.

- Search for the information – use a Web search engine to try to locate the information directly. The idea here is that the participant issued a search to try to directly find the information. An example is searching for "Mish Mish phone number" in a Web search engine to try to find the phone number for the local art store Mish Mish. The difference between searching for a resource and searching for the information can be subtle.

o   Example #3: If the task is to find the Web page of Dr. Stasko at Georgia Tech, then if the participant searched Google for "Dr. Stasko Georgia Tech", then it would be coded as "SE information" since the participant was using the SE to locate the specific information asked for in the task.

S1 to S2 comparison
This measure compares the starting strategy and resource used for each participant in session 1 to the starting strategy and resource used in session 2. These are tabulated with a code that indicates first the starting strategy comparison, then a slash character "/", and then a code that indicates the comparison of the resource used. The starting strategy comparison codes indicate if the starting strategy in session 2 was the same (S), different (X), or started from a mid-point in the session 1 finding, indicating a "learned" (L) waypoint or URL. The resource used comparison codes indicate if the resource used in session 2 was the same (S) or different (X) from the resource used in session 1. Table 41 indicates the combinations.

Table 41. Session 1 to Session 2 Comparison Codes

S/S	Same starting strategy on both days Same final resource used on both days
L/S	S2 starting strategy was a waypoint or URL apparently learned in S1 Same final resource used on both days
S/X	Same starting strategy on both days Different final resource used in S2 than in S1
L/X	S2 starting strategy was a waypoint or URL apparently learned in S1 Different final resource used in S2 than in S1
X/S	Different starting strategy used in S2 than in S1 Same final resource used on both days
X/X	Different starting strategy used in S2 than in S1 Different final resource used in S2 than in S1

5.7.1   FT1 – Dictionary look up

On both days, dictionary.com was used overwhelmingly for this task, followed by m-w.com. All the participants on both days used one of these two dictionaries as their resources, and almost all knew the URL to go to the resource directly. On both days, all participants who attempted the task successfully completed it. One participant in session 2 used MSN to search for www.dictionary.com, but for some reason ended up using m-w.com to get the definition. No participants tried to use a Web search engine to find the definition – a feature that Web search engines have recently started to support.

Session 1 n = 17			Session 2 n = 17			S1 to S2 n = 17	
Attempted	17		Attempted	17			
Completed successfully	17		Completed successfully	17			
			Not attempt or success in S1	17			
Resources			**Resources**			**Strategy diff**	
dictionary.com	14		dictionary.com	13		S/S	16
m-w.com	3		m-w.com	4		L/S	0
						S/X	0
						X/S	0
Strategy			**Strategy**			X/X	1
Go directly to start resource	16		Go directly to start resource	15		L/X	0
Search Web for information	0		Search Web for information	0			
Search Web for resource	1		Search Web for resource	2			

5.7.2 FT3 – Mish Mish art supply store phone number

Most participants on both days used some type of online phone directory such as yellowpages.com for this task. Two other strategies were, 1) to use a Web search engine to try to find information about the Mish Mish store, and 2) to look in the Blacksburg Electronic Village pages. Most participants knew how to go directly to some type of online phone directory and when presented with a task to find a phone number, they chose to use that resource. A point to note about this task is that for some reason the URL logging program had a problem with occasionally not including the URL yellowpages.com (the front page to the yellow pages site) in the log even though a participant visited this URL.

Session 1 n = 17			Session 2 n = 17			S1 to S2 n = 17	
Attempted	17		Attempted	17			
Completed successfully	17		Completed successfully	17			
			Not attempt or success in S1	17			
Resources			**Resources**			**Strategy diff**	
on-line yellow pages	11		on-line yellow pages	12		S/S	9
search engine results	3		search engine results	3		L/S	3
www.bev.net	3		www.bev.net	2		S/X	3
						X/S	0
Strategy			**Strategy**			X/X	2
Go directly to start resource	9		Go directly to start resource	12		L/X	0
Search Web for information	6		Search Web for information	3			
Search Web for resource	2		Search Web for resource	2			

5.7.3   FT6 – Find instructions about how to solve the Rubik's Cube

The choice of strategy for this task in both sessions was overwhelmingly to use a search engine to search for "rubiks cube solution" or something similar. A few participants appeared to guess a URL "rubikscube.com" that worked. An interesting aspect of this task is that while almost all the participants used the same starting strategy in both sessions, that many of them ended up at different resources in the second session than they had found in the first. This illustrates two things: first is that refinding can sometimes involve refinding the exact same information from the exact same resource, but sometimes using the same source does not matter as long as the information is found. The second point is that motivating refinding tasks in a laboratory setting can be challenging. This task could have required participants to refind exactly the same resource as they had found in the first session, but that might have required motivating the task more by asking the participants to read or study things about the page so that they would have a background on which to understand how/why to refind the particular resource (this was done with the newspaper article task in which participants were asked to read a news article). Since the participants in this study were only asked to find a page with a Rubik's cube solution, the particular page did not matter to them on either session. In a real-life setting, such a task would likely have been motivated by trying to actually solve a Rubik's cube and once a person had invested time learning a solution from one Web site, their refinding need might be based on re-locating the same Web site.

In this task, on both days, most participants used a search engine to locate a source of information about the Rubik's cube and its solution. Two interesting points: 1) the URLs for the resources are not very easy to remember in contrast to searching for "rubiks cube", 2) many people searched on both days but ended up at a different resource on day 2, and 3) the resources used on day 2 were more consolidated, perhaps indicating that people were more effective in their selection of resources on day 2.

Participants did not seem to "learn" a better path to the target information in this task. Most participants simply used a Web search engine again to search for "rubiks cube solution" or something similar. Several reasons may exist for this: 1) this strategy worked the first time and is straightforward and easy to remember and execute, 2) the URLs for the resources are not very easy to remember in contrast to issuing a search for "rubiks cube solution". This may be an example of least cognitive effort – it is easier to remember how to get to the information by remembering a search string than by remembering a URL. This is also an example of how the addressibility of information [Ram05] may be to remember a path rather than to learn a direct route.

Session 1 n = 17		Session 2 n = 16		S1 to S2 n = 16	
Attempted	17	Attempted	16		
Completed successfully	17	Completed successfully	16		
		Not attempt or success in S1			
Resources		**Resources**		**Strategy diff**	
lar5.com/cube	4	lar5.com/cube	5	S/S	5
jjorg.chem.unc.edu	2	jjorg.chem.unc.edu	6	L/S	0
jeays.net/rubiks.htm	3	jeays.net/rubiks.htm	3	S/X	10
other	8	other	2	X/S	0
				X/X	1
				L/X	0
Strategy		**Strategy**			
Go directly to start resource	3	Go directly to start resource	1		
Search Web for information	14	Search Web for information	15		
Search Web for resource	0	Search Web for resource	0		

5.7.4   FT10 – Atlanta Falcons score

On both days, ESPN.com was overwhelmingly used as the main source to find the football score. Almost all the participants that used ESPN.com went directly to the site by typing the URL. Only two participants used a different starting strategy in Session 2 than they had in Session 1 for this task. Four participants used a search engine to find a resource in Session 1. Of these, two of them selected atlantafalcons.com from the search results and remembered this URL, going directly to it in Session 2. Interestingly, the two participants that used a different starting strategy for Session 2 switched from going directly to ESPN.com to using a search engine. A point to note about this task is that for some reason the URL logging program had a problem with occasionally not including the URL espn.com (the front page to the site) in the log even though a participant visited this URL. Other URLs on the espn.com site appear to be logged correctly.

Session 1 n = 17			Session 2 n = 16			S1 to S2 n = 16	
Attempted	17		Attempted	16			
Completed successfully	17		Completed successfully	16			
			Not attempt or success in S1	16			
Resources			**Resources**			**Strategy diff**	
espn.com	14		espn.com	12		S/S	11
atlantafalcons.com	2		atlantafalcons.com	3		L/S	2
sportsillustrated.com	1		sportsillustrated.com	0		S/X	1
other	0		other	1		X/S	1
						X/X	1
						L/X	0
Strategy			**Strategy**				
Go directly to start resource	13		Go directly to start resource	12			
Search Web for information	1		Search Web for information	0			
Search Web for resource	3		Search Web for resource	4			

5.7.5   FT11 – Movie showtimes

In this task, participants looked for the starting time of a movie that they would like to see. Most participants knew a resource for this information and navigated directly to it. Given that all the participants were college students, it is not surprising that they were familiar with how to find information about movies. Most participants (11 of 13) used exactly the same strategy during the second session as they did in the first, indicating that participants had an established way to accomplish this task. Several resources were popular, including movies.yahoo.com, regalcinemas.com, hollywood.com, and movietickets.com.

Session 1 n = 14			Session 2 n = 13			S1 to S2 n = 13	
Attempted	14		Attempted	16			
Completed successfully	14		Completed successfully	16			
			Not attempt or success in S1	13			
Resources			**Resources**			**Strategy diff**	
movies.yahoo.com	6		movies.yahoo.com	6		S/S	11
hollywood.com	3		hollywood.com	2		L/S	0
regalcinemas.com	2		regalcinemas.com	2		S/X	1
movietickets.com	2		movietickets.com	1		X/S	0
other	1		other	2		X/X	1
						L/X	0
Strategy			**Strategy**				
Go directly to start resource	13		Go directly to start resource	11			
Search Web for information	0		Search Web for information	0			
Search Web for resource	1		Search Web for resource	2			

5.7.6   FT12 – Microsoft stock price

A number of different resources were used for this task, especially in the first session. Interesting points about the second session are: 1) 13 of the 17 participants used a similar or learned starting strategy, 2) no participants used a SE to find a resource. On each day, only a handful of participants were proficient and efficient with this task. For example, in session one only two participants appeared to know the ticker symbol "MSFT" for Microsoft stock. However, many participants knew some financial related Web site that they used. Some of these were unexpected, but quite logical. For example, one participant went to the Washington Post Web site in both sessions (possibly because they knew that stock quotes are available in printed newspapers). Similarly, one participant found the information at quicken.com in both sessions.

Examples:
        Go directly to a starting resource = go to Yahoo and click on finance link
        Search Web for information = search Google for "Microsoft stock"
        Search Web for a resource = search Google for "stock quote"

Session 1 n = 17			Session 2 n = 17 / 17			S1 to S2	
Resources			**Resources**			**Strategy diff**	
moneycentral.msn.com	4		moneycentral.msn.com	6		S/S	8
money.cnn.com	3		money.cnn.com	2		L/S	2
finance.yahoo.com	2		finance.yahoo.com	5		S/X	2
finance.lycos.com	2		finance.lycos.com	0		X/S	0
financial.washingtonpost	2		financial.washingtonpost	2		X/X	4
other	4		other	2		L/X	1
Strategy			**Strategy**				
Go directly to start resource	10		Go directly to start resource	13			
Search Web for information	4		Search Web for information	4			
Search Web for resource	3		Search Web for resource	0			

5.7.7   FT13 – Professor's Web site at Georgia Tech

This task illustrates an interesting set of finding behaviors. First, the strategy used by almost all the participants (12 of 15) was to use a Web search engine to locate the main Web site for Georgia Tech, even though the specific name "Dr. Stasko" was included in the task description. Participants were all students at Virginia Tech, which has a "people finder" feature on its main Web page. Thus, Virginia Tech students are accustomed to locating people at the university by going to www.vt.edu and searching for the person's name in the people finder. It appears that this mental model of how to find people at a university was strong enough to outweigh using a Web search engine to try to directly locate Dr. Stasko's Web site. One participant even searched for www.gt.edu, apparently guessing that since Virginia Tech is www.vt.edu, that Georgia Tech might be www.gt.edu. At the time of the study, Georgia Tech did not have a people finder feature on their home page, but it did have two links that attracted participants: a "faculty & staff" link, and a "site search" link. Many participants tried to use the site search to locate Dr. Stasko's Web page. However, the participants often had difficulty finding the correct page in the top search results. For example, the top three results of a site search for "Dr. Stasko" returned links to: 1) "CS7001 Project 1 – Anthropomorphic Interface Agents" – a class project, 2) "GVU Center: People: John Stasko" – a page about Dr. Stasko, and 3) "Webpage" – a page with a list of courses taught by various faculty members. Several participants were attracted to the "Webpage" and "CS7001" links, but did not try the "GVU Center" link, which would have taken them to Dr. Stasko's Web page. Several participants tried following the link for "Faculty & Staff" which was available on the Georgia Tech home page and also as a drop-down menu available from any of the Georgia Tech main pages. This link took participants to a part of the Web site for faculty, not a list of faculty. Despite these difficulties, 10 of 15 participants during the second session continued to use a general search engine to locate the Georgia Tech home page and then find a way to Dr. Stasko's Web site. This lends support to the idea that people may first decide if they can think of a source to find the information and if so, they then try to locate that source. As an indication of how strong patterns can be, one participant who used Google to search for Georgia Tech made the exact same spelling mistake in their first search on both days, they searched for "georgie tech".

Session 1 n = 13		Session 2 n = 13			S1 to S2 n = 13	
Attempted	15	Attempted	16			
Completed successfully	13	Completed successfully	16			
		Not attempt or success in S1	13			
Resources		**Resources**			**Strategy diff**	
gvu.gatech/~john.stasko	3	gvu.gatech/~john.stasko	6		S/S	7
cc.gatech/~john.stasko	8	cc.gatech/~john.stasko	3		L/S	0
gvu.gatech/gvu/people	2	gvu.gatech/gvu/people	1		S/X	5
cc.gatech/gvu/people	0	cc.gatech/gvu/people	3		X/S	0
	0		0		X/X	0
					L/X	1
Strategy		**Strategy**				
Go directly to start resource	0	Go directly to start resource	0			
Search Web for information	3	Search Web for information	4			
Search Web for resource	10	Search Web for resource	9			

5.7.8   FT14 Session 1 – Weather in Madrid, Spain

Most participants went directly to weather.com to accomplish this task. All 14 of the participants that went directly to a resource went directly to either weather.com or to the weather pages on Yahoo. The 3 participants that used a Web search engine found the information either on wunderground.com or on intellicast.com.

Session 1		
n = 17		
Attempted	17	
Completed successfully	17	
Resources		
weather.com	12	
weather.yahoo.com	2	
wunderground.com	2	
intellicast.com	1	
Strategy		
Go directly to start resource	14	
Search Web for information	2	
Search Web for resource	1	

5.7.9   FT14 Session 2 – Web site for The Record Exchange

For this task, almost all the participants eventually got to the Web page from a Web search engine (15 of 16). However, not all participants started with a Web search engine. Four participants tried going directly to a resource such as the VT home page, an on-line yellow pages, or guessing the URL. Only one of these four were able to find a link to the Web page from their starting resource – the participant who started at the VT home page went directly to the BEV pages and found a link for "The Record Exchange" that led them to "www.penduluminc.com/mm/welcome.htm" which appeared to be a music magazine that might somehow be related to The Record Exchange and may have been part of its Web site at some point.

Session 2		
n = 17		
Attempted	17	
Completed successfully	16	
Resources		
SE result → trexonline.com	15	
BEV → penduluminc/MM	1	
Strategy		
Go directly to start resource	4	
Search Web for information	12	
Search Web for resource	0	

5.7.10   FT15 – Panasonic VCR instruction manual

Task 15 was one of the most difficult tasks. In the first session, only 5 of 15 participants who attempted the task were verified to have completed it successfully[10]. In the second session, only 9 of 15 participants reported that they had completed it successfully. The task was difficult because Web search engines did not have the manual indexed from within the Panasonic Web site and the local search on the Panasonic site also did not return a link to the manual. If a participant looked for the manual by going to the Panasonic site, they had to notice a key link on the customer service page to get to the manuals. If they noticed this link, the task was fairly straightforward. However, many participants did not get to the key link or click on it. There was one successful participant who found the manual elsewhere on the Web by looking a page or two into the results of a Web search engine. It is interesting that more participants found the manual in the second session – they may have taken advantage of things they remembered not working in the first session. Also, many of the participants in the first session may have become frustrated at not finding it and starting fresh on the second day may have helped them locate the information.

Session 1 n = 5		Session 2 n = 5		S1 to S2 n = 5	
Attempted	15	Attempted	16		
Completed successfully	5	Completed successfully	9		
		Not attempt or success in S1	5		
Resources		**Resources**		**Strategy diff**	
panasonic.com	4	panasonic.com	4	S/S	3
mediacollege.com	1	mediacollege.com	1	L/S	2
				S/X	0
				X/S	0
Strategy		**Strategy**		X/X	0
Go directly to start resource	3	Go directly to start resource	4	L/X	0
Search Web for information	2	Search Web for information	1		
Search Web for resource	0	Search Web for resource	0		

5.7.11   FT16 – Weather in Blacksburg, VA

Task 16 was one of the easiest tasks. All the participants that attempted it successfully completed it and all went directly to a known source to find the information. The most popular source was weather.com (12 of 15 participants). Participants clearly had strong patterns for this task – there was no difference among any of the participants in the starting strategy used for the task from session 1 to session 2.

Session 1 n = 15		Session 2 n = 15		S1 to S2 n = 15	
Attempted	15	Attempted	17		
Completed successfully	15	Completed successfully	17		
		Not attempt or success in S1	15		
Resources		**Resources**		**Strategy diff**	
weather.com	12	weather.com	12	S/S	15
weather.yahoo	2	weather.yahoo	2	L/S	0
noaa.gov	1	noaa.gov	1	S/X	0
				X/S	0
Strategy		**Strategy**		X/X	0
Go directly to start resource	15	Go directly to start resource	15	L/X	0
Search Web for information	0	Search Web for information	0		
Search Web for resource	0	Search Web for resource	0		

---

[10] One additional participant indicated that they had found the VCR manual, but was not included in this analysis because it was unclear if they had found it.

5.7.12   FT17 – Academic calendar

Almost all participants during both sessions started this task by going to the main Virginia Tech Web page at www.vt.edu. Only three started with some other URL: one went to the main page for the College of Business at VT, one used a Web search engine to find the VT home page, and the other went directly to www.calendar.vt.edu, a site with a calendar of VT events. The information requested was located in several different locations on the VT Web site. Some participants used the same starting strategy in Session 2 as they did in Session 1, but ended up finding the information in a different location during Session 2. Participants seemed to require the most steps (URLs) when they tried to find the information within the www.calendar.vt.edu site. This allowed participants to page through an event calendar by month and display a summary of activities in either a day, week, or month view (the default was one day summary). Thus, participants needed to page through the calendar (i.e. load many URLs) to find the month and week in which they expected the activity to be located. Participants that looked for the information in the academics section were able to find in a listing of important academic dates. It is interesting to note that in Session 2, one participant thought they found the correct information, but in fact had located the dates from the previous year's academic calendar. Also, some participants, after going to the events calendar (www.calendar.vt.edu), took advantage of a link on the events calendar to the academic calendar.

Session 1 n = 14			Session 2 n = 14			S1 to S2 n = 14	
Attempted	14		Attempted	15			
Completed successfully	14		Completed successfully	15			
			Not attempt or success in S1	14			
Resources			**Resources**			**Strategy diff**	
calendar.vt.edu/main.php	7		calendar.vt.edu/main.php	5		S/S	8
academics/…/20032004	4		academics/…/20032004	5		L/S	0
academics	2		academics	1		S/X	4
currentstudents	1		currentstudents	2		X/S	2
			academics/…/20022003	1		X/X	0
						L/X	0
Strategy			**Strategy**				
Go directly to start resource	14		Go directly to start resource	13			
Search Web for information	0		Search Web for information	0			
Search Web for resource	0		Search Web for resource	1			

Chapter 6

Discussion

This chapter summarizes the main findings of this research and presents a discussion of the results for each research question. A preliminary model of search engine use in finding and refinding based on frequency and familiarity is also presented.

6.1 RQ 1 – Differences between finding and refinding

RQ1 – Do finding and refinding differ? Specifically, are there measurable differences between finding and refinding in terms of user knowledge/familiarity, perceived difficulty, search engine use, and task length?

The contributions of my work show that finding and refinding have differences for some types of tasks, but that for other tasks, no significant differences were observed. Task familiarity ratings were generally higher for refinding tasks and the perceived difficulty of a task can have effects on finding and refinding behaviors. No significant change in Web search engine use was observed between finding and refinding. These results have implications for the design of refinding tools, Web search engines, and additional studies of refinding. Tools and search engines can use these findings to help identify when a person is trying to refind information and to provide additional cues and integrate previously viewed results.

6.1.1 Main findings

This section summarizes the main findings for RQ1.

6.1.1.1 *Finding and refinding do have differences when defined by Session, but not for all tasks.*

Session had a main effect at $p<=0.01$ on task completion time (session 2 faster), number of URLs (session 2 lower), task familiarity (session 2 higher), before-task difficulty (session 2 rated easier), and after-task difficulty (session 2 rated easier). The Session x Task interaction was also significant at $p<=0.01$ for task completion time, task familiarity, and before-task difficulty. Post-hoc tests showed the specific tasks that differed from session 1 to session 2 – these are shown in Table 42.

Table 42. Tasks that were significantly different between the two sessions

Task completion time: (all faster in session 2)	Rubik's cube, professor Web site, and flight reservation
Task familiarity: (all more familiar in session 2)	Rubik's cube, professor Web site, grocery phone, art store phone, football score, stock price
Before-task difficulty: (all rated easier in session 2)	Rubik's cube, professor Web site, grocery phone, art store phone, stock price

Five tasks were not observed to be significantly different from each other in the Session x Task interaction tests: dictionary, movie times, local temperatures, academic calendar, and map distance. These tasks may not have differed from each other because participants already had high familiarity with many of them.

The lack of an effect of Session on task frequency was expected and is a reassuring result (see note about task frequency in section 4.4.4.1). It was not expected that in the course of the one week between session 1 and session 2 that participants would have changed the frequency with which they did any of the tasks. Not finding a significant effect of Session on Frequency is therefore a good check on the self-reporting methods used in this study and suggests that the participants answered the questions honestly and consistently.

6.1.1.2   *Finding and refinding do have differences when defined by Prior Frequency, but not for all tasks.*

High frequency tasks were completed more quickly, took fewer URLs, and involved less use of search engines than low frequency tasks. This indicates a difference between finding (low frequency tasks) and refinding (high frequency tasks). One session of experience improved the measures of task completion time (marginally significant increase, $p=0.08$), task familiarity ($p<=0.01$), and before-task difficulty ($p<=0.01$) from the session 1 measures.

6.1.1.3   *Search engine use was not observed to change from the first session to the second session.*

No significant main effect of Session or interaction effect of Session x Task was observed for either the number of uses of search engines, or for the ratio of the use of search engines to the total number of pages viewed. The lack of an observed effect of Session on search engine use is evidence that participants did not change their information seeking strategy for a given task from session 1 to session 2 (at least in terms of how they used search engines). While the lack of an observed effect does not guarantee that the effect is not present, this result is an interesting indication that users may use similar strategies to refind information as they did to find it. This result also addresses a fundamental issue in the study of information refinding – do users prefer to use search engines to refind information, or do they try to retrace their previous steps? The results here indicate that users may have patterns of access that are used over and over. Additional studies need to be conducted on larger samples of users to draw strong conclusions about this point. An observation about "known resources" is made based on the data from this study – if people know a reliable resource, they will be likely to try it as a starting strategy rather than using a general Web search engine. Many uses of on-line counterparts to paper resources such as telephone books and atlases were observed.

6.1.1.4   *Most tasks were completed in a similar amount of time in both sessions.*

Looking at the post-hoc analysis of the Session x Task interaction shows several important points. First, the mean task completion time was slightly lower for most tasks in the second session, but not significantly so. Most of the EP-15 tasks on both days took on average 50 to 100 seconds to complete and many did not differ significantly in terms of mean task completion time. This can be understood by considering that most of the EP-15 tasks were directed information finding tasks that asked participants to locate one specific item such as a football score or phone number. The two tasks that took significantly more time than the others were the task 18 (map distance) and task 5 (flight reservation). Each of these tasks required multiple steps so it is not surprising that they took more time than the other tasks.

Only three tasks had significantly different mean task completion times from session 1 to session 2: Rubik's cube (task 6), professor Web site (task 13), and flight reservation (task 5). One possible reason that the task completion time did not differ for most tasks from one session to the other is that participants were already familiar with many of these tasks such as looking up movie times, dictionary words, and local temperature information. In these cases, there may have been a ceiling effect – participants had already reached a high level of proficiency with these tasks before participating in the study and thus we observed a consistent level of performance in both sessions. This illustrates a limitation of using task completion time as a dependent measure – perhaps users would like to find the weather faster, but have already reached the minimum time using current technologies. Future studies should include user satisfaction and affective measurements to better understand users' happiness with and wants for refinding tools.

6.1.1.5   Participants perceived changes in familiarity, even though task completion times did not improve.

A group of six tasks differed significantly in terms of task familiarity ratings across the two sessions: Rubik's cube (task 6), stock price (task 12), professor Web site (task 13), football score (task 10), art store phone (task 3), and grocery phone (task 8). All of these six except for the football score task also had significantly different before-task difficulty ratings from session 1 to session 2. These tasks appear to be tasks that participants either gained a lot of familiarity with from the first session, or for which they initially overestimated the difficulty. It is important to note that these six tasks were among the seven tasks with the lowest familiarity ratings in session 1 – thus, these were the tasks that had the most room for improvement in terms of familiarity. Tasks which participants had high prior familiarity in the first session did not benefit significantly from the experience gained in the first session.

6.1.2   Interesting observations

In this section, interesting observations about findings for RQ1 are discussed.

6.1.2.1   Task familiarity for unfamiliar tasks was influenced by only one session of experience.

Two tasks are interesting to explore in relation to perceived task familiarity. First, consider the Rubik's cube task (task 6). In one sense, this can be a fairly easy task once a user realizes that typing "Rubik's cube solution" into a Web search engine will return many relevant Web pages. However, this task had the lowest mean familiarity and the most difficult mean before-task difficulty rating of all the EP-15 tasks. While it is not surprising that participants reported that they were not familiar with the location of a Web page with instructions on how to solve the Rubik's cube, it is interesting that they thought that this would be a difficult task. In the session 2 ratings, after participants realized how easy it was, the mean task familiarity and mean before-task difficulty ratings were similar to the other tasks. The mean task completion time for this task was lower for session 2 than session 1, indicating participants did actually improve their performance on this task. It is unclear exactly why participants initially thought this would be a difficult task. Almost all the participants used a search engine for this task, indicating that they knew a good strategy. Perhaps in relation to things like sports scores, phone numbers, and weather information this type of information seemed harder to locate, with the possibility that it might be hard to find.

Second, consider the flight reservation task. Interestingly, this task had low session 1 familiarity, but did not have a significantly higher familiarity for session 2. This may represent a task that participants had difficulty learning in just the one session and points out an interesting aspect of information refinding – that some tasks that seem difficult at first may be easy to remember how to do, but that other tasks may require more learning to become familiar with how to do them. This is an interesting area for future research – tasks that are easy for people to remember how to do them may not benefit as much from refinding tools as tasks for which people have trouble remembering them or for which they learn to perform them more slowly. It is likely that the tasks that were easier to learn had analogs that participants were already familiar with but did not realize it at first until they had performed the task once. For example, the Rubik's cube task can be a fairly easy task once a user realizes that by typing "Rubik's cube solution" into a Web search engine that many relevant results will be returned. However, in contrast, finding a flight reservation can be a complex, multi-stage task and may be perceived as more difficult when becoming familiar with how to do it.

6.1.2.2   Perceived difficulty and perceived familiarity are not always inversely related.

Consider a task that illustrates a point about the difference between the before-task difficulty measure and the task familiarity measure (also measured before the task). The map distance task (task 18) had a significantly higher mean task familiarity in session 2 than in session 1, but did not have a significantly higher before-task difficulty rating. It also had a significantly lower task completion time in session 2 than in session 1. This may indicate that participants better understood the task during session 2, but that they knew it to be a difficult task due to the complexity inherent in the task (it required several steps). This raises an interesting point about the complexity of tasks in finding and refinding: for some tasks, users may

realize that they will have to repeat all the steps to refind the desired information, but for other tasks, they may recognize that they can use a shortcut to "jump" back to particular point in their search process.

6.1.3   Implications

The contributions of my work show that finding and refinding have differences for some types of tasks, but that for other tasks, no significant differences were observed. By recognizing that some types of refinding are different from finding and that other types of refinding are similar to finding, tools, search engines, and PIM research can benefit. For example, my results suggest that for tasks that users are familiar with, that they are likely to have well-known patterns and techniques for re-accessing the information. Refinding tools may not need to help users with these tasks – they already know how to do them. This is an interesting point because several previous research projects attempted to help users create short-cuts to frequently accessed information. These results suggest that users may already have established paths for these types of information. A related question is, what triggers a user to change their access pattern for a piece of information? Perhaps as the value of the information access path becomes greater, or as the user realizes the importance of the information, they will take the time to create a shortcut or find a more efficient access pattern. Aspects of this question are explored further in the model of refinding behavior presented in section 6.5. Similarly, recognizing tasks that users are likely to have trouble refinding is an important part of developing effective refinding tools. My results showed a set of tasks for which users improved significantly from session 1 to session 2 – these tasks should be examined further to see if a longer gap between finding and refinding still has improved performance, or if users need help refinding these types of information.

Search engines can benefit from recognizing distinctions and similarities between finding tasks and refinding tasks. By recognizing that a person is attempting to refind information, search engines could provide additional cues to help users. Additionally, as part of the finding process, search engines could provide additional information and context that will help users in a later refinding process. For example, some search engines are starting to subdivide results based on the type of resources (e.g. Web pages, phone book, movies, etc.). Search engines could also recognize when users are re-issuing the same or similar query and highlight previously explored results (similar to the ideas explored in the ScentTrails work at Xerox PARC [OC03]), or allow users to save queries for reuse (A9.com offers a similar feature through their toolbar application).

The examples of how task familiarity for unfamiliar tasks was influenced based on one session of experience (see section 6.1.2) illustrate an important design point for refinding tools – a participant's *perception* of how easy or difficult a task will be may affect the strategies they use and decisions they make about if and how they can relocate specific information. The actual difficulty may be quite different from the user's perception, but if the user feels that there is a large cognitive distance between their task goal and the refinding mechanisms provided by an available tool, they may not attempt the refinding using that tool and might try a different approach or drop the task altogether.

Thus, refinding tools need to help users understand what tasks they can help with and accurately convey to users the types of resources and information available. One step in this direction is visibly displaying the set of databases over which a search will be conducted, such as is done on the A9.com Web site. On this site, a list of databases such as "Web", "books", "movies", "yellow pages" is presented with check boxes to allow the user to select which databases to include in the query. However, such interfaces still face the challenge of helping users map the myriad of possible user tasks onto what needs to be a relatively simple initial search interface.

6.2   RQ 2 – Task Differences

RQ2 – Does task and familiarity with a task affect finding and refinding? What types of tasks are easy and which are difficult? Specifically, how do task and task familiarity (knowledge of how to perform the task) affect user knowledge/familiarity, perceived difficulty, search engine use, and task length?

My work contributes to this research question by showing that task is an important factor for both finding and refinding tasks. Furthermore, not all tasks were different from finding to refinding (i.e. there were Session x Task interaction effects). These result are consistent with results from information science literature in which task was found to be an important factor to consider when studying information seeking behaviors [WRM+02] [TFK+03]. In my study, tasks were found to fall into three general categories: fairly *easy* tasks which participants were familiar with, *difficult* tasks that involved several steps or topics with which the participants were unfamiliar, and a "*medium*" set of tasks that were in-between the easy and difficult sets. Task affected the use of Web search engines and prior familiarity with a task also had significant effects.

6.2.1   Main findings

6.2.1.1   *The task being done does have an effect on finding and refinding, but some tasks were similar.*

Task had a significant main effect at $p<=0.01$ on all the dependent variables. However, as described in the summary for RQ 1 in section 6.1.1, there was also a significant Session x Task interaction for task completion time, task familiarity, and before-task difficulty at $p<=0.01$. Post-hoc tests showed that some tasks had similar dependent measures. One reason for this may be that many of the tasks included in this study were well defined, directed tasks to find a specific, single piece of information.

6.2.1.2   *Tasks tended to fall into three groups: easy, medium, and hard.*

Four tasks tended to have mean values for the dependent variables indicating that these tasks were easier than other tasks: task 1 (dictionary), task 11 (movie times), task 16 (local temperature), and task 17 (academic calendar). Four other tasks tended to have mean values for the dependent variables indicating that these tasks were harder than other tasks: task 5 (flight reservation), task 6 (Rubik's cube), task 13 (professor Web site), and task 18 (map distance). The remaining tasks tended to group together in the middle of the rankings of the mean values for the dependent variables.

Table 43. Tasks with low and high dependent measure values

	Low/Easy	**High/Difficult**
Task Completion Time	16, 1, 6, 3, 17	18, 5
Number of URLs	1, 10, 16, 6, 3, 11, 8	18
Number of searches	16, 17, 1, 11, 10, 8, 3, 5	13, 18
Search ratio	16, 17, 1, 5, 11, 8	18, 13, 6
Task Frequency	16	6, 12, 5
Inverse Task Familiarity	17, 11, 1, 16	6, 12, 13, 5
Before-task difficulty	11, 1, 16	5, 12, 18, 13, 6, 3
After-task difficulty	11, 17, 6, 1, 16	18, 5
Summary	1, 11, 16, 17	5, 6, 13, 18

6.2.1.3 *Task had a significant effect on search engine use.*

There was a main effect of Task on number of searches and search ratio. Task 6 (Rubik's cube), task 13 (professor Web site), and task 18 (map distance) had more search use than other tasks. These tasks were well suited for a search engine for two main reasons. First, the task description contained specific information that would be likely to lead to the information requested by the task. For example, "Rubik's cube" is a distinctive name that would be likely to return pages from a search engine about the Rubik's cube puzzle. In the task to find a professor's Web site, the name of the university (Georgia Tech) was included in the task description and was often used as a search engine query to locate a starting resource. In the map distance task, Web searches were often made to help find the address of the two endpoints.

6.2.1.4 *Familiarity had significant effects on finding and refinding.*

Prior familiarity had a significant main effect at $p <= 0.02$ on all the dependent variables. Tasks with high familiarity had lower task completion times, involved fewer URLs, fewer uses of Web search engines, and were perceived as easier in both before- and after-task difficulty ratings.

There was also a significant Session x Prior Familiarity interaction for task completion time, number of URLs, before-task difficulty, and after-task difficulty at $p <= 0.02$. In the Session x Prior Familiarity interactions, typically there was no gain from session 1 to session 2 for the high prior familiarity tasks, but there was improvement for the medium and low prior familiarity tasks.

6.2.1.5 *Familiarity with a task affected search engine use.*

Prior Familiarity had a main effect on both number of searches and search ratio at $p <= 0.01$. Post-hoc tests showed a significant difference among all three levels of Prior Familiarity for both number of searches and for search ratio at the $p <= 0.01$ level. No significant Session x Prior Familiarity interaction was observed for either number of searches or search ratio.

These are important results that show that prior familiarity has an effect on search engine use that does not appear to change from session 1 to session 2. These results suggest that search engine use is inversely proportional to prior familiarity for a continuum of tasks from familiar to unfamiliar.

There is still an open question about how familiarity changes over the long-term. In this study, many participants rated tasks in session 2 with higher familiarity measures than they rated them in session 1, but this familiarity increase was only for a one-week time gap between finding and refinding. This short-term increased familiarity did not appear to affect search engine use (recall that prior familiarity – familiarity as measured in session 1 – was used to measure the effect of familiarity on search engine use). One possibility is that if a familiarity with a task was maintained over a longer period of time (e.g. two months) that the search engine use might also change. However, if the task was not performed very often and remained unfamiliar, then search engine use might not decrease for the task. This would be an interesting aspect of refinding to explore in a future study.

In this study, familiarity was measured by the question "How well do you know how to get to the information?" There may be some tasks that for which people always chose to use a search engine, even if they do the task fairly often. It is a valid question to wonder how a person would rate their familiarity with where the information is located for such a tasks. Since they always use a search engine to find it, they might say that they don't really know where it is (i.e. they don't know the absolute address of the information – the URL). However, they might say that they do know where the information is, or at least that they know "how to get to the information" since they have an *a priori* plan that they use to arrive at it – their plan just happens to include the use of a search engine. A user following this plan probably searches for the same keyword string each time and may even know the location of the target in the search results page. This subtle distinction in the wording of the question "Do you know where the information is located?" versus "Do you know how to get to the information?" could have an effect on the responses in the situations outlined above and would be an interesting topic to address in a future study.

6.2.2   Interesting observations

Some observations about task differences and similarities:

- All participants knew a dictionary URL
- Most participants knew an on-line yellow pages URL
- Most participants knew a movie listing Website
- Few participants knew how to find a stock price.
- Many participants were not familiar with making flight reservations – possibly because all the participants were students at Virginia Tech and many Virginia Tech students are from areas of Virginia within driving distance.

It also appears that people have fairly strong models of the Web and how particular types of Web sites work. For example, participants appeared to expect that university Web sites would have some type of "people finder" and that major retail stores would have some type of "store locator" feature on the corporate Web site.

While working on this study, six major ways to refind information on the Web became apparent:

1) remember the URL
2) remember a sequence of links
3) create a bookmark or shortcut
4) remember a specific string for a Web search engine
5) remember a starting point / waypoint
6) start from scratch and find it again (i.e. refinding = finding again)

6.2.3   Implications

My research shows that task is an important factor to consider not only for finding, but also for refinding. The results are consistent with prior research into information seeking and extend the importance of task to refinding. In addition, prior familiarity with a task is also an important factor that affects refinding behaviors.

These results have implications for the study of finding and refinding and also for the design of tools. Given the importance of task, it is important to consider taxonomies of tasks when studying finding and refinding. In my study, a set of 18 tasks were developed based on empirically gathered data about frequent and infrequent tasks done on the Web. A better understanding of what tasks are similar is needed, especially for the study of personal information management. Many PIM tasks are similar in that they are well defined and typically focus on locating a specific piece or pieces of information. However, which of these types of tasks are related and along what factors? My research begins to shed light on this question (see Table 43 in Section 6.2.1.2), but more work is needed to more fully understand the space of PIM tasks. Existing taxonomies such as [BJW+99] [MPC01] [Bro02] are good starting points for exploration.

Tools for finding and refinding are likely to benefit by recognizing and considering the type of task being performed. For example, tasks that are complex or require several steps may benefit more from shortcuts and tool assistance than easier tasks. Tasks that are perceived to be difficult also may benefit from additional cues and subdivisions of information presentation. Tasks that involve searches could benefit from features to save and re-issue previous searches and to highlight the results that were explored.

One session of experience often led to large increases in perceived familiarity. In the Session x Prior Familiarity interactions, typically there was no gain from session 1 to session 2 for the high prior familiarity tasks, but there was improvement for the medium and low prior familiarity tasks. This result indicates that that one session of practice appears to increase participants' performance on the tasks – in some cases to the level of a participant with high prior familiarity. Practice effects are well-known aspects

of HCI studies and are an important aspect to consider both in the study of refinding and for implementing tools to help with refinding.

In this study, tasks were mostly directed tasks to find a single piece of information and may have been easy to gain knowledge of in just one session. Time is also an important factor. The two sessions for each participant in this study were scheduled about one week apart. These results suggest that for a directed refinding task to find a single piece of information within one week of finding that people are able to make good use of their memory of the finding experience. Future work should examine how this would change if the time between sessions were much greater (e.g. several months).

6.3   RQ 3 – Finding to Refinding Relationship

RQ3 – How does the relationship of the refinding task to the finding task affect the refinding process? Specifically, does it affect user knowledge/familiarity, perceived difficulty, search engine use, and task length?

My work makes several significant contributions to this area. The study of refinding as a distinct process is new and the distinctions of exact, path, and subset are contributions to the study of refinding. Results of my research indicate that refinding moved or deleted items can be so difficult that users may abandon the attempt, or not try at all. Refinding an item from a subset task took longer and was perceived as more difficult and is an area that should be studied further. My results on tasks to refind information on the same path may have been affected by an unintentional bias in the task grouping that was not discovered until the analysis stage. However, these "path" tasks are still of interest and may orient toward resources that contain sets of valuable information and thus may be more familiar and require less use of Web search engines.

6.3.1   Main findings

6.3.1.1   *Refinding moved/deleted items is very difficult and may be impossible in some cases.*

Only two of 17 participants were able to refind the news story they had found in the first session. This indicates that tasks where information is moved or deleted from the location that it was found can be very difficult for users.

6.3.1.2   *Task similarity does appear to affect refinding, but the results for path tasks are unclear and more studies of task similarity are needed*

In the analysis of exact, path, and subset tasks in session 2, Task similarity had a significant main effect at $p<=0.01$ on search ratio, task familiarity, before-task difficulty, and after-task difficulty and at $p<=0.05$ on task completion time and number of searches. The differences observed fell into two main groups: differences of the path tasks on search engine use, and differences in which subset tasks were more difficult than exact or path tasks. As discussed further in the section below, the differences of the path tasks in terms of search engine use may have been caused by unintentional bias. However, it is also possible that path tasks tend to be oriented toward resources that have sets of information and are more familiar to users and thus have less use of Web search engines. Path tasks should be examined further in future studies. The differences between the subset tasks and the exact and path tasks suggest that refinding an item from a set may pose special challenges and should be studied further.

6.3.1.3   *Potential task bias – path tasks had much lower use of search than exact tasks*

An interesting result in both analyses of task similarity was that number of searches and search ratio were both significantly lower for path tasks than for exact tasks. This is interesting because the exact and path

tasks were not expected to be very different in terms of their dependent measures. Path tasks differed only in that they requested slightly different information in the session 2 task than in the session 1 task. For example, a session 1-path task might be to find the *phone number* for a business, and the session 2 task would be to refind the *address* for the same business.

In addition, in the analysis of exact and path tasks across the two sessions, path tasks were rated as more familiar and were perceived as easier in the before-task ratings, but took longer. These differences between exact and path tasks in terms of task completion time, task familiarity, and before-task difficulty were not observed in the analysis of the EPS-15 session 2 tasks in section 5.5.2.

When trying to understand these results, a potential bias was found that was unintentionally introduced when the tasks were created. The path tasks were all tasks that the author could easily think of two similar pieces of information that could be found on the same information path. For example, the Virginia Tech academic calendar included dates for *spring break* and for the *last day of classes*. By creating the tasks this way, these tasks also share a common property that was unintended – the author was able to easily think of a source with two similar pieces of information. It is likely that participants were also easily able to think of these information sources and this would explain why these tasks have lower search engine use than the other tasks and were rated as more familiar. However, as noted earlier, it may be that path tasks tend to be oriented toward resources that have sets of information and are more familiar to users.

This potential bias affects the analysis of task similarity. The differences between exact and path and between subset and path may be due to this bias rather than due to the presence of an actual effect. An implication of this is that experimenters must be very careful when designing tasks for finding and refinding. This possible bias was introduced in a completely unintentional way and was not realized until the data analysis stage. However, the comparisons between exact and subset tasks provide interesting results that indicate that subset tasks were more difficult than exact tasks. Additional studies should further examine task similarity to investigate possible effects of exact, path, subset, and moved tasks.

6.3.1.4   *Refinding subset tasks were perceived as more difficult and took longer than exact tasks.*

In the analysis of task similarity for session 2, the subset tasks had a significantly longer mean task completion time than the exact tasks ($p<=0.05$). The subset tasks also had significantly lower familiarity and more difficult before- and after-task difficulty ratings ($p<=0.01$) than the exact tasks. These results indicate that refinding an item from a set of similar items (as in the subset tasks) can be more difficult and is perceived as more difficult than refinding a single item (as in the exact tasks).

There are several important points to note about this result. Only two subset tasks were given to participants and each of the subset tasks involved finding two items in session 1 and then refinding one of the two in session 2. *Participants were allowed to select* which of the two items to refind in session 2. Based on participant comments, many appear to have selected the item for which they had the strongest memory. Refinding a particular item from a larger set of found, similar items is expected to be an even more challenging problem. For example, people often browse the Web looking for information about a topic or find information about a purchasing decision. As they browse, they see many pieces of similar information located on many different Web pages and Web sites. Trying to refind a single page or item some time after such a browsing session can be very challenging. Additional research on refinding subsets of information is needed.

6.3.2   Interesting observations

One of the subset tasks was to find two sweatshirts that the participant would be interested in buying as a gift for a friend. As part of the task, participants were asked to indicate the name of each sweatshirt and why they had selected it. This information was requested to help engage the participants in the task. The participants in this study were all undergraduates at Virginia Tech and many appeared to have considerable domain knowledge of shopping on-line for clothing. Many participants knew the URLs of Web sites for particular brands of clothing and commented that they frequently made purchases from these sites and/or knew them to be of good quality. This task also illustrated some domain terminology that the author was

not aware of prior to the study, but that many of the participants seemed well versed in – hooded sweatshirts are commonly referred to as "hoodies". This type of domain knowledge can be particularly important in finding tasks. Differences between domain experts and novices are an important area of research in information science [Wil04] [HS00] [JCW03].

6.3.3   Implications

Next to refinding moved information, refinding a subset of information when a larger set of related information was found may be one of the most difficult types of refinding tasks for many people.

Refinding an item or items that were found along with a set of many other similar items was perceived to be more difficult and took longer than other tasks. These types of tasks that involve refinding a subset of similar items that were found appear to be among the most challenging refinding tasks. Future studies of refinding should consider these types of tasks as a area to be examined in more depth.

Tools for finding and refinding are likely to benefit by recognizing and considering the similarity of a task from finding to refinding. For example, tasks that involve refinding a subset of information that was seen previously could benefit from special interfaces to allow users to quickly scan sets of Web pages and to highlight distinguishing features among the items in the set. Tasks that involve information that has moved or that has been deleted from the Web could benefit from tools that either automatically archive certain types of information, or that make it easy for users to save selected information. One idea would be to incorporate a button on Web browsers that allows users to turn on and off "recording" of the Web pages that they view.

Constructing tasks for use in controlled studies is a difficult task and care must be taken not to introduce biases. Even though the tasks in my study were developed with the help of an empirical study of frequent and infrequent tasks done on the Web, a potential bias exists when considering the path tasks as a group. This bias may be a natural characteristic of the grouping, or may have been introduced unintentionally when the tasks were created.

6.4   RQ 4 – Task Details and Strategies

RQ4 – What strategies are used in finding and refinding? Are there observable patterns of strategy use across types of tasks or from finding to refinding for a single task?

The contributions of my work indicate that users do have patterns of access that are followed for many types of tasks and that users are well versed in a (small) set of Internet resources (i.e. they know the URL for these sites and can type them directly into a Web browser to teleport to the site). Many of these known resources are Internet counterparts of longstanding paper resources such as dictionaries, phone books, newspapers, and maps. In addition, particpants often used the same starting strategies to refind a piece of information that they did to find it. In addition, there were cases in which participants used shortcuts or waypoints they appeared to have learned from the finding process to help simplify the refinding. In addition, often users ended their refinding at the same resource they used when finding. However, in some cases, the location of the particular resource used did not matter and the information was "refound" from a different resource than it was found. This is an example of the fungibility of information.

6.4.1   Main findings

6.4.1.1   *Users have a set of known Web sites, many of are Internet counterparts of common real-world resources*

For some tasks, few significant differences were found among them in terms of search engine use: task 1, task 5, task 8, task 11, task 16, and task 17. For all these tasks, participants appeared to have a known Web site that they knew would contain the information being requested, and thus, did not make much use of search engines. A summary of these tasks, the traditional resource and the Internet resource are shown in Table 44 below.

Table 44. Traditional and Internet Resources

Task Number	Task Description	Traditional (Pre-Internet) Resource	Internet Resource
1	Dictionary	Dictionary	dictionary.com, m-w.com
5	Flight reservation	Travel agent	on-line travel agencies
8	Grocery store phone	Yellow pages	on-line yellow pages (e.g. yellowpages.com, smartpages.com)
11	Movie times	Newspaper – Entertainment section	Portal/Movies Web site (e.g. movies.yahoo.com)
16	Local temperature	Newspaper / Phone line	weather.com
17	Academic calendar	Academic handbook	vt.edu Web site

Each of these tasks has a strong pre-Internet connection to an information source. These well-known, sources have been put into electronic form on the Internet and people appear to still "go to the resource" when looking for the corresponding types of information.

6.4.2   Interesting observations

When trying to refind information about tourist attractions in San Francisco, one participant had a particularly interesting interaction sequence. First, they remembered a specific thing to search for – "gold rush trail tours san fransico [sic]". They tried it, but did not spell "San Francisco" correctly. The participant was audibly disappointed that the specific tour that she had remembered to search for did not show up. Instead of clicking the "Did you mean" link at the top of the Yahoo search results, she searched again, this time just for "san fransico [sic]", still visibly struggling with the spelling, even though the "Did you mean" link showed the correct spelling. In typing the second search string, she first typed, "san franciso [sic]", and then "corrected" herself with "san fransico [sic]". This incorrect spelling "fransico" may have been reinforced because the top search result from the previous search had the words "san fransico [sic]" highlighted (because even though her first search was misspelled, it matched Web pages that also had the same misspelling). When the results of the second search came up, she realized this was still not correct and was audibly frustrated. At this point, she started to try to correct the spelling, and noticed the "Did you mean" link with "San Francisco" and clicked on it. From this results page, instead of trying to re-formulate her original query about the gold rush trail, she clicked a Yahoo link to a city guide for San Francisco and arrived at a page of San Francisco attractions. This page had a link and information about Alcatraz Island, another attraction she had viewed on the first day. On the task information form, she filled in information about Alcatraz instead of continuing to search for the gold rush trail.

This is an example of several things:

- Spelling affecting search and retrieval
- Aids presented by the search engine failing to attract the user (the "Did you mean" aid) until after several attempts and a reformulation of the original query
- The difficulties in motivating an information need in controlled studies.

It appears that the participant chose a different attraction than the one she originally thought of because she happened to find a different one sooner. Perhaps she even abandoned looking for the gold rush trail after the initial search failed to find it and decided to try for a more general approach (i.e. "San Francisco") just to refind *something* that she had found during the first day session. Similar issues were reported by Fidel et al. [FDD+99] in a study of high-school students who were given assignments to use the Web to search for information for about plants. The investigators observed that in some cases, "it was not unusual for them [the students] to change their minds about the selected plant if they thought they could not find the required information about the chosen one" [FDD+99, p. 28].

Another interesting interaction sequence occurred for the Mish Mish phone number task. In this case, the participant went to Google and searched for "Mish Mish Art Supply Store". Many results came back, but none for Mish Mish specifically. At this point, the participant changed strategies and typed in the URL "yellowpages.com", searched there, and found the phone number. However, in the first Google search results page, several of the summary descriptions on the results page had text indicating the name and phone number (i.e. "Store Name, Blacksburg, VA (540) 555-1234"). This is an example where it appears that the user fixed on the idea of finding a particular type of resource, such as the Web site for Mish Mish, and did not even notice the presence of the desired information on the search engine results page. Jakob Nielsen [Nie06] has commented on how the inclusion of information in search engine result summaries can sometimes reduce users "clicking through" to go to the Web site containing the information since it is displayed in the search results summary, although that was not the case in this example.

6.4.3   Implications

People tend to follow similar patterns of access for many refinding tasks as they did when they originally found the information. In addition, people have well-established methods for accessing particular types of information.

Tools to support refinding might benefit by recognizing patterns of access. Several existing tools have explored this, but mainly to help create shortcuts for frequently visited sites. For example, several projects have attempted to help point out connections between documents in a user's personal information space (e.g. RememberanceAgent [RS96], MarginNotes [Rho00] and Haystack [AKS99]). While these are potentially useful, many users are likely to already have established patterns of access for information that they re-access frequently. They may get more benefit from having refinding tools recognize when they are recreating an infrequent pattern of access that may indicate a refinding attempt.

The observation that people have a set of known resources that they refer to on the Web for particular types of information has important implications for finding and refinding. In my study, participants were observed teleporting directly to on-line dictionaries, yellow pages, movie timetables, weather information services, and sources of academic information. These results indicate that people maintain the notion of going to a resource on the Web, even when search engines are available. Recently, search engines have begun integrating additional resources and "searching" over multiple resources for a single query. For example, search engine A9.com allows users to search over the Web, movies, yellow pages, and many others sources and then presents the results in newspaper-style columns, separated by resource type. As search engines evolve, it will be interesting to see if users' strategies and approaches to finding and refinding information from these known resources will change. For example, a recent Pew Internet and American Life study [RS05] noted that there has been a increase in the use of local search features that search engines are offering (e.g. a query that includes a zip code will present localized results for that geographic region).

6.5   Model of How Frequency and Familiarity Affect Finding and Refinding

Several major results of this work helped motivate a preliminary model of search engine use in finding and refinding:

- Frequency and familiarity affect search engine use
- No effect of Session was observed on search engine use
- Tasks fell into three groups: easy, medium, and hard

These results suggested a model of finding and refinding in which a task may move through general stages based on: 1) the level of familiarity with how to locate the information in the task, and 2) the level of frequency with performing the task. The model is illustrated in Figure 40. The line represents a possible movement through the stages. Aspects of this model are also described in [CP05b]
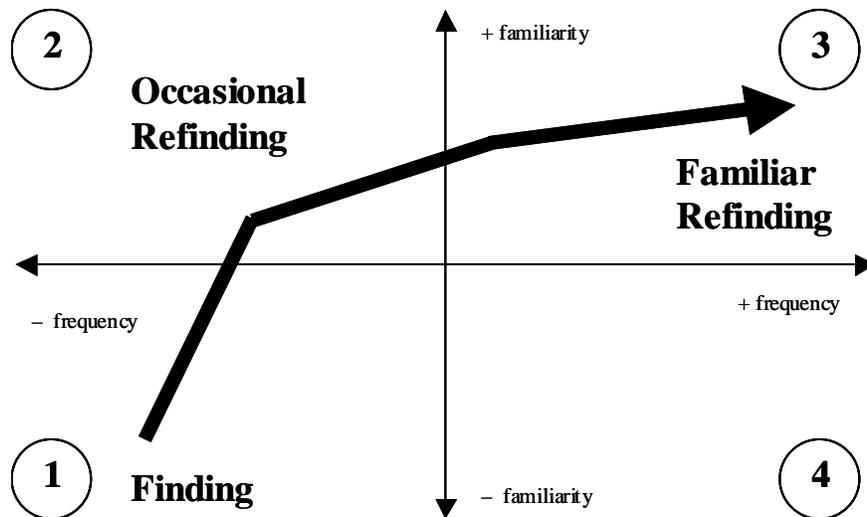
Figure 40. Model of How Frequency and Familiarity Affect Finding and Refinding

Quadrant 1 – Low frequency, low familiarity
In quadrant 1, tasks are either finding tasks or are performed with low enough frequency that they are similar to finding tasks (i.e. refinding = finding again). Users may not know the future value of the information for these tasks, or the nature of the task may be that it is a rarely performed task.

Quadrant 2 – Low frequency, high familiarity
Tasks in quadrant 2 are tasks that the user has learned something about, but that they still perform only occasionally. Many of these tasks may be done more frequently than the ones in quadrant 1, but still would not be considered high frequency. For example, a student may register for classes once every semester – a low frequency task – but they are likely to have become familiar with the process of finding the classes, registering, etc.

Quadrant 3 – High frequency, high familiarity
Quadrant 3 tasks are ones that have become frequent, well-known tasks. Users may have set up shortcuts or bookmarks for these tasks, or may have taken the time to learn specific access paths such as the URL or a series of links that will lead them to the information. Somewhere between quadrant 2 and 3 is likely to be

where users recognize the importance of the information/task and may decide to invest time in creating or learning a faster access path.

Quadrant 4 – High frequency, low familiarity
There are probably only a few tasks in quadrant 4, but it is included in the model for completeness. Quadrant 4 tasks are difficult by nature, no matter that a user may perform them somewhat frequently. Perhaps the user has not invested time to create or learn a shortcut, or perhaps some aspect of the task is dynamic, making it difficult to predict. For example, the task of looking for yesterday's top news stories is by nature difficult because many of the news stories will have moved. Frequently accessing a PDF file that does not have a link on a Web site, but requires a user to type a long, tricky URL could be considered a quadrant 4 task (assuming that the user did not invest the time to create a shortcut).

This model of refinding is an early-stage model that is suggested by several results from my research. Future work should test this model on a larger sample size and explore the features of each quadrant in more detail, especially to investigate what motivates users to move a task from one quadrant to another and how their access patterns change and what triggers them to invest time to create and learn shortcuts.

Chapter 7

Conclusions and Future Work

In this chapter, a summary of main findings is presented, contributions of this dissertation are outlined, and implications are discussed. At the end of the chapter, areas of future research work are described.

7.1   Summary of Main Findings

Brief summaries of the main findings for each research question are given in the following pargraphs.

RQ1 – Do finding and refinding differ?
Finding and refinding did have differences, especially for tasks that were perceived to be difficult. Search engine use was not observed to be affected by session, but was affected by task. Familiarity with a task was influenced by finding the information in the first session, but did not necessarily reduce the amount of time it took to refind the information.

RQ2 – Do task and familiarity with a task affect finding and refinding? What is easy/difficult?
Task and familiarity with a task both had effects on task completion time and search engine use. Tasks that were more familiar had less search engine use. Tasks that were generally easier included: looking up the definition of a word (dictionary), movie times, weather information, and academic calendar information. Tasks that were generally more difficult included: looking up airline flight information, locating a solution for the Rubik's cube, finding a professor's Web site, and finding the distance between two locations (map).

RQ3 – How does the similarity of the task from finding to refinding affect the refinding process?
The similarity of a task from finding to refinding was found to affect refinding. Refinding for subset tasks was perceived as more difficult and took longer than refinding for exact tasks. Refinding when the information moved from the location at which it was first found can be very difficult.

RQ4 – What strategies are used in finding and refinding?
Many users have a set of known Web sites that are often used as starting points for finding and refinding. Many of these sites are on-line counterparts of paper resources such as dictionaries, telephone directories, and newspapers. For many tasks, participants used the same starting strategy (i.e. they went to the same starting Web site) when refinding as they did when finding, indicating strong patterns of access.

7.2   Contributions and Implications

In this dissertation, I have made five main contributions to the field of information refinding and personal information management.

1. Identified, defined, and explored dimensions of information refinding.
2. Measured differences between finding and refinding in terms of factors and effects
3. Identified sets of easy and difficult finding and refinding tasks
4. Found patterns of Web information access
5. Developed a preliminary model of information refinding

In this section, each of these contributions is briefly summarized and implications for information design, tool features, and the study of refinding are discussed.

7.2.1   Dimensions of Refinding

When I started this work, there was little existing groundwork in the area of information refinding and the area was not well defined. One of the main contributions of my research is that I have helped identify, define, and explore the main dimensions of refinding and why they are important. Much of the discussion in Chapter 2 outlines important concepts and factors for refinding and relates prior study of them in related fields to their role in refinding. The importance of frequency, familiarity, and perceived difficulty are all explored in this dissertation. While many of these concepts are not new, their importance and application to refinding is a part of the contribution of this work.

Two of these concepts (frequency and familiarity; and task similarity) are discussed in more detail in the following paragraphs.

7.2.1.1   *Frequency and Familiarity*

To the best of my knowledge, my work is the first to examine the effects of frequency and familiarity on refinding. In a strict sense, information refinding by an individual can be viewed dichotomously – the information being sought has either been seen before and the current task is refinding, or this is the first time the information is being looked for and the task is finding. In a more relativistic sense, refinding can be considered in terms of frequency – how often does the searcher look for this information? For very frequent tasks, a user may have highly developed methods of access and may be more likely to "teleport" directly to the information. For medium frequency tasks, the refinding process may involve a mixture of recall and recognition. For very infrequent tasks, refinding may truly be "finding again".

7.2.1.2   *Task Similarity*

Information refinding can be a challenging problem when the information is static, but information on the Web is highly dynamic in terms of its location and lifetime. Studies have documented the dynamic nature of Web pages [Spi03] [BC00] [Koe02]. The ever-changing nature of the Web makes refinding even more difficult. Consider the task of refinding a news story that you had seen yesterday on the home page of an on-line news site. When you first saw the story there was probably a link to it on the front page, but how do you get to it a day later? Is there a search feature? If so, there may be many stories related to the topic – how would you know which one it is? What if the page has been deleted?

Task similarity refers to the similarity of a task from finding to refinding. Two related concepts are also introduced here: *information similarity* and *location similarity*.

> Task similarity (same)
> > Day 1: Find today's weather
> > Day 2: Find today's weather

Information similarity (same)
> Day 1: Find today's weather
> Day 2: Find the weather forecast you found yesterday

Location similarity (different location)
> Day 1: Find today's weather
> Day 2: Find the weather forecast you found yesterday
> (the information is not linked from the same place it was found)

When trying to refind information, several possibilities exist about the relationship of the original finding instance to the current refinding need. For example, a person may seek to refind information that was found along the path of the original finding task, but that was not the specific target information of the original finding task. I call this aspect of refinding *task similarity* and believe it is an important aspect of how people approach refinding tasks that has not be considered previously.

In this dissertation, I investigated four relationships, described below along with an example of each. Note that these are not an exhaustive set of all possible types of task, information, and location similarity, but rather they are a starting point for exploration of this important concept. In addition, there may be other dimensions of interest in addition to task, information, and location similarity.

- Exact – the information being sought when refinding is exactly the same as the information that was originally found.
  - Finding: phone number for the XYZ grocery store on Main St.
  - Refinding: phone number for the XYZ grocery store on Main St.

- Path – the information being sought when refinding is on the same path as the information that was originally found.
  - Finding: phone number for the XYZ grocery store on Main St.
  - Refinding: phone number for the XYZ grocery store on Oak Lane

- Subset – the information being sought when refinding is a subset of the information that was originally found.
  - Finding: three sweatshirts as a possible gift for a friend
  - Refinding: a specific one of the three sweatshirts originally found

- Moved – the information being sought when refinding has moved from the location where it was originally found.
  - Finding: a news story on the front page of an news Website
  - Refinding: the same news story (moved off the front page)

7.2.2   Differences in finding and refinding / Factors and Effects

Finding and refinding information on the Web are different processes with different characteristics.
Task type has a large effect on refinding.
Users' prior knowledge is an important factor in how refinding is accomplished.
In this dissertation, I identified significant differences in finding and refinding as measured by differences in task completion time, task familiarity, and perceived task difficulty. These differences depend greatly on the task being performed and the prior experiences of the user. Users have different experience and knowledge with different types of tasks and perceive different tasks as being different levels of difficulty. Prior familiarity with a task affected task completion time, number of URLs, and search engine use.

These results suggest that a "one-size-fits-all" approach may not work well for information refinding. Design strategies and tools that work well for one type of task may not be well suited for other types of

tasks. In addition, tools that support keyword searches of personal information may be helpful for many tasks, but users may prefer to directly apply their existing knowledge to help refind information and to remember context rather than use keyword search, as suggested in [BN95] and [BS04].

Refinding is not just finding again
These results are an important contribution not only because specific effects on measured variables were observed, but also in a larger sense, because there has been debate in the personal information management community about whether refinding is simply "finding again". At a recent workshop on PIM, this question of "refinding = finding again?" was a topic of discussion [JB05, p.16-20]. The results of this dissertation provide insights into the refinding process and illustrate specific cases where refinding was found to have differences from finding. An important implication of this result is that tools for refinding should recognize and leverage the differences between finding and refinding.

Factors that affect "the keeping decision"
The results of this research shed light on aspect of the problem of deciding whether to keep information or to rely on being able to refind it on the Web [Jon04, Conclusion section]. Different information is refound with differing degrees of ease. Some information can easily be refound on the Web, potentially affecting a user's decision not to "keep" this information in a personal storage system. However, other information is more difficult to return to, possibly influencing a user's decision to keep a copy of it. The results of this research show that factors such as the amount of knowledge that a user has about how to refind the information, and how difficult the task is perceived to be are important components in the finding and refinding process and are likely play an important role in the keeping decision.

7.2.3   Tasks that are easy and difficult

Identifying tasks that are easy and difficult for finding and refinding is important for the understanding and development of tools and Web sites that support people's refinding needs. By including a large sample of representative tasks in this study it has been possible to identify tasks that were generally found to be easy and tasks that were generally found to be difficult. Tasks that were easiest included finding a dictionary definition, weather information, phone numbers, and sports scores. For many of the "easy" tasks, the information could be found by going to a well-known resource such as an on-line dictionary or yellow pages. The most difficult tasks included finding map/route information, relocating a news article, looking for specific product manual, and finding flight information. The difficult tasks involved: several steps; moved, or difficult to locate information;, or a domain that was unfamiliar to the participants.

These differences illustrate that while some information is simple to get back to on the Web, other information is more difficult, pointing out specific areas for future research and tool development. One aspect of this is that tools are needed to help support users in refinding tasks that they find difficult. Several tools have focused on trying to make it easier for users to re-access information that they access frequently (i.e. by automatically recognizing patterns). While this may be useful, tools are also needed that support difficult refinding tasks. Several types of difficult refinding tasks are described below along with possible features/tools to help support them.

Complex tasks – Support for complex tasks should recognize the need to manage complexity *within* a particular Web site and the need to manage complexity *across* different Web sites. Managing information and complexity within a site might be best handled as part of the site design – examples include site tools to keep track of recent items viewed, to save and reissue recent searches, to make comparisons of items, and to mark or tag information on the site. Managing information and complexity across different sites is potentially more difficult and might be best handled as part of a Web browser or browser tool. There are many existing tools that could help with this – tabbed browsing and bookmarks are two examples – but these require knowledge and effort on the part of the user and can be difficult to share among different computers or browsing sessions. For example, for the task in my study to find the distance between two geographic locations many users found the addresses of the two locations and then copied these into the start and destination fields on a map directions Web site. Even with multiple windows, this would require

several switches back and forth to the different sites and several copy and paste operations to get all the information required. When trying to refind the information, refinding the addresses might be done more quickly, but copying and pasting them into the fields on the map Web site would likely not be any faster. This illustrates a situation where a "save search" feature on the map Web site could make refinding faster.

Moved information – Trying to refind information that has moved is very difficult for users and could be better supported by both Web sites providing information and by browsing tools. Web sites, especially ones that provide large amounts of dynamic information, could be designed for better refinding support. Consider a daily news Web site. Instead of providing a general search feature for the site that returns a flat list of results, the site could provide a feature to view the site as it looked on a particular day or range of days. This could be challenging for sites that change continuously, but could greatly help refinding. People may remember the location of the information on the page, or some other spatial or graphical cue that is lost if the information is presented differently than it was when seen the first time. Browser tools could help users by providing easy mechanisms to save pages (or sets of pages) of interest and then providing an easy-to-use search over those pages. Something as simple as a "save this page locally" button on the browser toolbar could help. However, such a feature should be easy to use and require minimal input by the user – deciding what to call the page or what folder to put it in could be a large enough barrier that a user might not use it (although these should be options).

Unfamiliar domains – Finding and refinding information in unfamiliar domains is a special challenge. Many Web sites attempt to provide help to users who are unfamiliar with the domain, but acquiring specialized domain knowledge can be very challenging as noted by Bhavnani [Bha02]. Refinding for these types of tasks could benefit from tools to help users save or log important pages and information found.

7.2.4   Patterns of Access

Often, participants used the same starting strategies (i.e. went to the same starting Web sites) when refinding as they did when finding. In addition, no significant effect of Session was observed for search engine use, but there was a significant effect of Task on search engine use. This suggests that task more than session affected the use of Web search engines to locate information in this study. For some tasks, users relied on a set of known resources rather than search engines. These results illustrate strong patterns of Web information access. Maglio and Barrett [MB97a] also observed strong patterns of Web information access.

Patterns of access can help users retrace familiar steps to refind information, but may also inhibit refinding in the form of "false memories" – steps may be remembered as part of a pattern of access that were not actually taken in a specific case [MB97a]. Refinding tools may need to help users factor out the common aspects of a search in order to help users recognize a particular path they are trying to refind. Summaries of browsing histories may help with this, particularly if they include graphical representations of the Web pages viewed. Several projects have investigated the use of thumbnail images in recognizing and organizing information [ATH+03] [WRM+02] [KGE02] [RCL+98]. These efforts should be continue to be explored and investigated as part of refinding processes.

7.2.5   Preliminary Model of Refinding

A preliminary model of refinding is presented in this dissertation (see section 6.5). The model is based on a user's prior frequency with a task and their familiarity with where the information is located. While this is an early model, it highlights the concept that access patterns may change based on factors such as the frequency of access, familiarity of location, and recognition of the value of the information. Understanding these factors is an important aspect of designing effective tools and features to support refinding.

There are certainly additional factors that affect access. For example, if a person is not at their primary computer, they may not have access to their shortcuts or bookmarks. This type of distributed cognition appears to plays an important role in information refinding – people often rely on information stored in

their computing environment (or on post-it notes attached to their monitor) to relocate information. Distributed information management tools such as on-line bookmarks address some of these aspects, but also require users to take the effort to input information and log in to the system. In addition, some people have privacy concerns about storing lots of personal information on Web servers. Additional factors that affect finding and refinding need to be explored in more detail and integrated into a more complete model of refinding.

7.2.6   Additional Implications and Discussion

Different tasks require different tools, designs, and features to support refinding. For example, users appear to be fairly good at refinding information nuggets and routinely accessed information. Tools may be better focused on helping users get back to information found through more complex steps and to provide help in refinding particular information from among a group of related pages/sites that were visited when finding.

A number of currently available tools are aimed at helping people refind personal information, including the recently released Google Desktop Search, Apple Spotlight, and A9 toolbar. Many "personal search" tools are geared toward searching files, documents, and email messages stored on a user's personal computer. Most provide some support for searching Web pages, but often the pages must have either been saved to the user's hard disk or still be in the Web browser's cache file (i.e. only recently viewed Web pages). In addition, several projects have developed add-ons or special versions of Web browsers that will automatically save or help facilitate saving Web pages that a user browses (examples include: TrailBlazer, Slogger, ScrapBook, and the University of Calgary's Internet Explorer History System). There are also services (such as Furl and Spurl) that allow users to store bookmarks and browsing history information on-line so that they are accessible from multiple computers and can be shared with other people.

These efforts to help support searches of personal information appear to be steps in the right direction, but are still in the early stages of evolution and may not address all the needs of users for refinding. While keyword-based search tools can be very powerful, several studies have suggested that keyword search is not always the method chosen by users to refind information, even when available [Jon04] [TAA+04]. The results of my study support this: for tasks that participants felt they had good knowledge of how to get to the information, they used search engines less relative to the total number of pages loaded. Waypoints and context can play a strong role, beyond just helping to relocate information. Context can play an important reminding function that may not be satisfied through typical search interfaces in which keywords are typed in and a flat list of matching results is presented. Interfaces that include a variety of search, exploration, and presentation techniques, including visualizations of information and paths traversed are interesting possibilities for future research.

7.3   Future Work

There are still many areas that need more exploration and understanding regarding information refinding and personal information management (PIM). In this section, I describe several broad areas for future work and then outline particular studies that would extend the field.

7.3.1   Broad Areas

Broadly, there are several areas that have not been explored in much detail within the umbrella of information refinding and PIM:

- Understanding Refinding – The main study described here reveals some insight into the processes, difficulties, and factors that affect how people attempt to refind information on the Web. However, there are still many important questions in this area. For example, how does the elapsed time between finding and refinding affect the refinding process? Can searchable histories and visual representations of browsing sessions help users refind information more effectively? How can refinding strategies be categorized, and how can tools help with these strategies? In addition, more exploration of tasks is needed, especially to understand the characteristics of tasks that are difficult for refinding. This is likely to require interviews, surveys, and/or field studies to observe refinding tasks that cause users difficulties.

- Mobile Issues – Little work to date has examined how people manage and refind information while they are in transit or away from one of their primary computers. This area of study includes issues related to information synchronization, the use of multiple devices, the similarity of interaction and presentation on different devices, and understanding users' mobile refinding needs.

- Personal Search – Tools that index a user's entire hard disk, cataloging files, emails, stored Web pages, and other information are now available. The impact of these tools on how people refind information and deal with PIM is yet to be seen. Several studies have indicated that users sometimes like to use keyword search, but sometimes like to relocate information manually. The possible impact of these new personal search tools on refinding and PIM behaviors should be investigated.

- Personal Information Management – In general, PIM has started to gain more attention as the research community begins to address the problems users face with "information overload". There are many areas of PIM that need additional exploration – files, email, Web pages, contact information, calendars, and growing areas such as photos and multimedia content. Increased interest in tagging and metadata impacts many of these areas and should be explored.

7.3.2   Specific Studies

In this section, several specific studies are proposed to further examine areas of information refinding and PIM.

7.3.2.1   *Analysis of Refinding Strategies*

I believe it could be valuable to look in more depth at the question of what strategies people use for finding and refinding various types of information. This would require a significant data collection effort and then the careful development of a taxonomy of strategies based on the data.

Ideas for strategies and behaviors to explore include:

- When are typed URLs used? What URLs do people know? How many URLs do people know?
- Does the use of the "back" button in Web browsers provide evidence of particular types of strategies?
- How can finding and refinding strategies be classified and categorized? Can a taxonomy of strategies currently used be created? A good starting point for this would be to start with task taxonomies and models of information seeking from information science.
- It may be very important to conduct some studies in this area that do not focus as much on measurements of time, number of URLs, etc. and instead investigate what users are thinking as they search. Verbal protocols such as "think-aloud" may provide informative insights.

7.3.2.2　　Additional Controlled Studies

Controlled experiments could be conducted to examine additional factors of interest. One such study would be to examine the effects of the elapsed time between finding and refinding. A study to do this could involve three (or more) sessions:

- An initial session to find information on the Web
- A second session scheduled 1-2 weeks later to refind some of the information and possibly to find some additional information
- A third session scheduled 1-2 months later to refind information from both sessions 1 and 2.

Comparisons between the strategies used in all three of the sessions could be made to see which tasks had consistent patterns and for which tasks the refinding strategies changed. In addition to specific, directed tasks, it would be interesting to include several less well defined tasks in which the participants would find information by browsing for pages and Web sites about the topic and then refind specific information.

Additional controlled studies could examine nuances of task familiarity. For example, there may be some tasks that for which people always chose to use a search engine, even if they do the task fairly often. It is a valid question to wonder how a person would rate their familiarity with where the information is located for this type of task. Since they always use a search engine to find it, they might say that they don't really know where it is (i.e. they don't know the absolute address of the information – the URL). However, they might say that they do know where the information is, or at least they know "how to get to the information" since they have an *a priori* plan that they use to arrive at it – their plan includes the use of a search engine, they probably search for the same keyword string each time, and they may even know the location of the target in the search results page. This is related to the notion of the addressibility of information as described by Naren Ramakrishnan [Ram05]. A subtle distinction in the wording of the question "Do you know where the information is located?" versus "Do you know how to get to the information?" could have an effect on the responses in the situations outlined above and would be an interesting topic to examine.

7.3.2.3　　Personal Search Tools

Google, Apple, and Microsoft are each developing tools to help people search over their own personal information. It would be interesting to conduct studies to see which types of tasks may benefit from these types of personal search tools. These studies could be combined with several of the factors that were identified as important in this dissertation.

7.3.2.4　　Affective and Satisfaction Measures of Finding and Refinding

A limitation of using task completion time, number of URLs and other objective dependent measures is that they may not capture all the important aspects of an interaction. Perhaps users would like to find a particular piece of information faster, but have already reached a minimum time using current technologies. Perhaps users often are able to find and refind the information they seek, but are frustrated by the processes supported by existing tools. Future studies should investigate user satisfaction and affective measurements to better understand users' wants and needs for refinding tools.

7.3.2.5   Expanded Version of Frequent/Infrequent Survey

Understanding the tasks that people do frequently and infrequently is an important aspect of understanding refinding. An expanded version of the preliminary survey presented in this dissertation with a larger sample size and additional questions and methods could help give insights into difficult tasks. The expanded version of the survey could include 100-200 participants from several different groups (e.g. students, knowledge workers, less experienced Web users, seniors).

The survey could have the following basic format and could be conducted using an on-line survey administration tool:

> Background information
> Collect information about tasks
> > Ask for 5 frequent tasks
> > > User describes the task
> > > User rates frequency
> > Ask for 5 infrequent tasks
> > > User describes the task
> > > User rates frequency
> Recall steps for the tasks
> > Randomly ask for 3 tasks of each type
> > > Describe the steps you use to do this task in as much detail as you can recall.
> > Randomly ask for 3 tasks of each type
> > > Do you have any shortcuts/bookmarks for this task?
> Ask about difficult refinding tasks
> > Describe two situations where you had trouble/difficulty refinding some information you had seen on the Internet recently. Include enough detail for us to visualize the situation.

The coding for the data could be as follows:

> Have two raters categorize tasks as before
> > Compare categories and frequencies to Pew Internet study results
> Have two raters code each description of tasks
> > (random presentation, blind to whether frequent or infrequent)
> > as either orienteering, teleporting, or mix of both
> Have two coders categorize refinding problems

Ideas for additional analysis include:

> Do frequent tasks involve more teleporting?
> Do frequent tasks have more shortcuts/bookmarks?
> What types of tasks do people have trouble refinding?

This page intentionally left blank.

References

[Abi97]   Abiteboul, S. (1997). Querying Semi-Structured Data. In F. N. Afrati and P. Kolaitis (Eds.), Proceedings of the 6th International Conference on Database Theory (ICDT-97), Delphi, Greece, January 8-10, 1997. (pp. 1-18). Springer.

[ABC98]   Abrams, D., Baecker, R. and Chignell, M. (1998). Information Archiving with Bookmarks: Personal Web Space Construction and Organization. In Conference Proceedings on Human Factors in Computing Systems. (pp. 41-48). Los Angeles, CA, USA. 1998. ACM Press/Addison-Wesley Publishing Co.

[AKS99]   Adar, E., Karger, D. and Stein, L. A. (1999). Haystack: Per-User Information Environments. In Proceedings of the Eighth International Conference on Information Knowledge Management. (pp. 413-422). ACM.

[ATH+03]   Amento, B., Terveen, L., Hill, W., Hix, D., and Schulman, R. (2003). Experiments in Social Data Mining: The TopicShop System. ACM Transactions on Computer-Human Interaction 10(1):54-85.

[And80]   Anderson, J. R. (1980). Cognitive Psychology and Its Implications. San Francisco: W.H. Freeman.

[AJK05]   Aula, A., Jhaveri, N., and Käki, M. 2005. Information search and re-access strategies of experienced Web users. In Proceedings of the 14th international Conference on World Wide Web (Chiba, Japan, May 10 - 14, 2005). WWW '05. ACM Press, New York, NY, 583-592.

[BN95]   Barreau, D. and Nardi, B. A. 1995. Finding and reminding: file organization from the desktop. SIGCHI Bull. 27, 3 (Jul. 1995), 39-43.

[BMK97]   Barrett, R., Maglio, P. P. and Kellem, D. C. (1997). How to Personalize the Web. In Conference Proceedings on Human Factors in Computing Systems. (pp. 75-82). 1997. ACM Press.

[BCS+95]   Belkin, N.J., Cool, C., Stein, A. & Thiel, U. (1995). Cases, scripts and information-seeking strategies: On the design of interactive information retrieval systems. Expert Systems with Applications, 9(3):379-395.

[BMC93]   Belkin, N.J., Marchetti, P.G., and Cool, C. (1993). BRAQUE: Design of an interface to support user interaction in information retrieval. Information Processing and Management, 29(3):325-344.

[BDH+03]   Bellotti, V., Ducheneaut, N., Howard, M., and Smith, I. 2003. Taking email to task: the design and evaluation of a task management centered email tool. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Ft. Lauderdale, Florida, USA, April 05 - 10, 2003). CHI '03. ACM Press, New York, NY, 345-352.

[Bha02]   Bhavnani, S. K. 2002. Domain-specific search strategies for the effective retrieval of healthcare and shopping information. In CHI '02 Extended Abstracts on Human Factors in Computing Systems (Minneapolis, Minnesota, USA, April 20 - 25, 2002). CHI '02. ACM Press, New York, NY, 610-611.

[BB02]   Bhavnani, S., and Bates, M. 2002. Separating the knowledge layers: Cognitive analysis of search knowledge through hierarchical goal decompositions. Proceedings of the American Society for Information Science and Technology 39(1):204-213.

[BBJ+03]   Bhavnani, S. K., Christopher, B. K., Johnson, T. M., Little, R. J., Peck, F. A., Schwartz, J. L., and Strecher, V. J. 2003. Strategy hubs: next-generation domain portals with search procedures. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Ft. Lauderdale, Florida, USA, April 05 - 10, 2003). CHI '03. ACM Press, New York, NY, 393-400.

[BS04]   Boardman, R. and Sasse, M. A. 2004. "Stuff goes into the computer and doesn't come out": a cross-tool study of personal information management. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM Press, New York, NY, 583-590.

[BM98]   Bonura, T. and Miller, J. R. (1998). Drop Zones: An Extension to LiveDoc. ACM SIGCHI Bulletin. 30: 59-63.

[BC00]   Brewington, B., and Cybenko, G. (2000). How dynamic is the Web? Computer Networks, 33:257-276.

[Bro02]   Broder, A. (2002). A Taxonomy of Web Search. ACM SIGIR Forum, 36(2):3-10.

[BBG+01]   Brush, A. J. B., Bargeron, D., Gupta, A. and Cadiz, J. J. (2001). Robust Annotation Positioning in Digital Documents. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. (pp. 285-292). Seattle, WA, USA. 2001. ACM Press.

[BH00]   Budzik, J. and Hammond, K. J. (2000). User Interactions with Everyday Applications as Context for Just-in-time Information Access. In International Conference on Intelligent User Interfaces 2000. (pp. 44-51). ACM.

[BJW+99]   Byrne, M., John, B., Wehrle, N., Crow, D. (1999). The Tangled Web We Wove: A Taskonomy of WWW Use. In Proceedings of the 1999 Conference on Human Factors in Computing Systems. ACM Press, New York, 544-551.

[Bys02]   Byström, K. (2002). Information and Information Sources in Tasks of Varying Complexity. Journal of the American Society for Information Science and Technology, 53(7):581-591.

[BJ95]   Byström, K. and Järvelin, K. (1995). Task Complexity Affects Information Seeking and Use. Information Procesing and Management, 31(2):191-213.

[CGG00]   Cadiz, J. J., Gupta, A. and Grudin, J. (2000). Using Web Annotations for Asynchronous Collaboration Around Documents. In Proceedings of the ACM 2000 Conference on Computer Supported Cooperative Work. (pp. 309-318). Philadelphia, PA, USA. 2000. ACM Press.

[CPP04]   Capra, R., Pinney, M., and Pérez-Quiñones, M.A. Refinding Is Not Finding Again. Technical Report TR 05-10. Department of Computer Science, Virginia Tech.

[CP03]   Capra, R. and Pérez-Quiñones, M.A. (2003). Re-Finding Found Things: An Exploratory Study of How Users Re-Find Information. Technical Report cs.HC/0402001, Computing Research Repository (CoRR).

[CP05a]   Capra, R., and Pérez-Quiñones, M.A. (2005). Mobile Refinding of Web Information Using a Voice Interface: An Exploratory Study. In Proceedings of the 2nd Latin American Conference on Human-Computer Interaction (CLIHC 2005), October 2005, Cuernavaca, Mexico.

[CP05b]   Capra, R., and Pérez-Quiñones, M.A. (2005). Using Web Search Engines to Find and Refind Information. *IEEE Computer 38*(10): 36-42.

[CPR01]   Capra, R., Pérez-Quiñones, M. and Ramakrishnan, N. (2001). WebContext: Remote Access to Shared Context. In Workshop on Perceptual User Interfaces. Orlando, FL, USA. November 2001.

[CMN83]   Card, S.K., Moran, T.P., and Newell, A. (1983). The Psychology of Human-Computer Interaction. Lawrence Erlbaum, Hillsdale, NJ.

[CDT00]   Choo, C.W., Detlor, B., and Turnbull, D. (2000). Information Seeking on the Web: An Integrated Model of Browsing and Searching. First Monday, 5(2). http://www.firstmonday.org/issues/issue5_2/choo/index.html

[CGJ+03]   Cockburn, A., Greenberg, S., Jones, S., Mckenzie, B. and Moyle, M. (2003) Improving Web Page Revisitation: Analysis, Design And Evaluation, IT & Society, 1(3):159-183.

[Cot02]   Cothey, V. (2002). A Longitudinal Study of World Wide Web Users' Information-Searching Behavior. Journal of the American Society for Information Science and Technology, 53(2):67-78.

[DV00]   Denoue, L. and Vignollet, L. (2000). An annotation tool for Web browsers and its applications to information retrieval. In Recherche d'Information Assistée par Ordinateur (RIAO2000).

[DV01]   Denoue, L. and Vignollet, L. (2001). Personal Information Organization using Web Annotations. In WebNet 2001.

[DAW98]   Dey, A. K., Abowd, G. D. and Wood, A. (1998). CyberDesk: A Framework for Providing Self-Integrating Context-Aware Services. In Proceedings of the 1998 International Conference on Intelligent User Interfaces. (pp. 47-54). San Francisco, California, USA. ACM Press.

[DB01]   Ducheneaut, N. and Bellotti, V. 2001. E-mail as habitat: an exploration of embedded personal information management. *interactions 8*, 5 (Sep. 2001), 30-38.

[DCC+03]   Dumais, S. T., Cutrell, E., Cadiz, J. J., Jancke G., Sarin, R., and Robbins, D. C. (2003). Stuff I've Seen: A System for Personal Information Retrieval and Re-Use. Proceedings of SIGIR 2003.

[DFL+88]   Dumais, S. T., Furnas, G. W., Landauer, T. K., Deerwester, S. and Harshman, R. (1988). Using Latent Semantic Analysis to Improve Access to Textual Information. In Conference on Human Factors in Computing Systems (CHI '88). (pp. 281-285). Washington, D.C., USA. 1988. ACM Press.

[EFH+76]   Eckstrom, R.B., French, J.W., Harman, H.H. and Derman, D. (1976). Kit of factor-referenced cognitive tests. Educational Testing Service, Princeton, NJ.

[EH89]   Ellis, H., and Hunt, R. (1989). Fundamentals of Human Memory and Cognition, Fourth Edition. Dubuque. Wm. C. Brown.

[ES91]   Erickson, T. and Salomon, G. (1991). Designing a Desktop Information System: Observations and Issues. In Human Factors in Computing Systems Conference Proceedings on Reaching Through Technology (CHI '91). (pp. 49-54). 1991. ACM Press.

[FBC+02]   Farrell, S., Buchmann, V., Campbell, C. S. and Maglio, P. P. (2002). Information Programming for Personal User Interfaces. In Proceedings of the 2002 International Conference on Intelligent User Interfaces. (pp. 190-191). 2002. ACM Press.

[FFG96]   Fertig, S., Freeman, E., and Gelernter, D. (1996). "Finding and reminding" reconsidered. *SIGCHI Bulletin, 28*(1):66-69.

[FDD+99]   Fidel, R., Davies, R.K., Douglass, M.H., Holder, J.K., Hopkins, C.J., Kushner, E.J., Miyagishima, B.K., and Toney, C.D. (1999). A Visit to the Information Mall: Web Searching Behavior of High School Students. *Journal of the American Society for Information Science, 50*(1):24-37.

[FMM01]   Ford, N., Miller, D., and Moss, N. (2001). The Role of Individual Differences in Internet Searching: An Empirical Study. *Journal of the American Society for Information Science and Technology, 52*(12):1049-1066.

[FWF+02]   Ford, N., Wilson, T.D., Foster, A., Ellis, D. and Spink, A. 2002. Information Seeking and Mediated Searching, Part 4, Cognitive Styles in Information Seeking. *Journal of the American Society for Information Science and Technology, 53*(9):728-735.

[GPS99]   Golovchinsky, G., Price, M. N. and Schilit, B. N. (1999). From Reading to Retrieval: Freeform Ink Annotations as Queries. In Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. (pp. 19-25). Berkeley, CA, USA. ACM Press.

[Gwi04]   Gwizdka, J. 2004. Email task management styles: the cleaners and the keepers. In CHI '04 Extended Abstracts on Human Factors in Computing Systems (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM Press, New York, NY, 1235-1238.

[GVU98]   GVU Center at Georgia Tech. (1998). GVU's 10th WWW User Survey (October 1998). Retrieved March 19, 2002, from http://www.gvu.gatech.edu/user_surveys/

[Har02]   Hargittai, E. (2002). Beyond Logs and Surveys: In-Depth Measures of People's Web Use Skills. Journal of the American Society for Information Science and Technology, 53(14):1239-1244.

[HLO99]   Heck, R. M., Luebke, S. M. and Obermark, C. H. (1999). A Survey of Web Annotation Systems. Retrieved April 4, 2002, from <http://www.math.grin.edu/~rebelsky/Blazers/Annotations/Summer1999/Papers/survey_paper.html>

[HS00]   Hölscher, C. and Strube, G. (2000). Web Search Behavior of Internet Experts and Newbies. *Computer Networks 33*, 337-346.

[Ima02]   iMarkup Solutions. (2002). iMarkup Homepage. Retrieved March 13, 2002, from <http://www.imarkup.com/>.

[JP01]   Jansen, B., and Pooch, U. (2001). A Review of Web Searching Studies and a Framework for Future Research. Journal of the American Society for Information Science and Technology 52(3):235-246.

[JS06]   Jansen, B. and Spink, A. (2006). How are we searching the World Wide Web? A Comparison of nine search engine transaction logs. Information Processing and Management 42(1):248-263.

[JCW03]   Jenkins, C. Corritore, C., Wiedenbeck, S. (2003). Patterns of Information Seeking on the Web: A Qualitative Study of Domain Expertise and Web Expertise. IT&Society, 1(3):64-89.

[JK96]   John, B., and Kieras, D. (1996). Using GOMS for User Interface Design and Evaluations: Which Technique? ACM Transactions on Computer-Human Interaction, 3(4):287-319.

[Jon86]   Jones, W. P. (1986). The Memory Extender Personal Filing System. In Conference Proceedings on Human Factors in Computing Systems. (pp. 298-305). Boston, MA, USA. ACM Press.

[Jon04]   Jones, W. (2004). Finders, Keepers? The Present and Future Perfect in Support of Personal Information Management. First Monday 9(3). March 1, 2004.

[JB05]   Jones, W., and Bruce, H. (2005). A Report on the NSF-Sponsored Workshop on Personal Information Management. January 27-29, 2005, Seattle, WA. Accessed on February 26, 2006 from http://pim.ischool.washington.edu/final%20PIM%20report.pdf

[JBD01]   Jones, W., Bruce, H. and Dumais, S. (2001). Keeping Found Things Found on the Web. In Proceedings of the Tenth International Conference on Conference on Information and Knowledge Management. ACM Press, New York, 119-126.

[JBD03]   Jones, W., Bruce, H. and Dumais, S. (2003). How Do People Get Back to Information on the Web? How Can They Do It Better? In Proceedings of the 9th IFIP TC13 International Conference on Human-Computer Interaction (INTERACT 2003), Zurich, Switzerland, September 1-5, 2003.

[JD86]   Jones, W. P. and Dumais, S. T. (1986). The Spatial Metaphor for User Interfaces: Experimental Tests of Reference by Location versus Name. ACM Transactions on Office Information Systems 4(1): 42-63.

[JF97]   Jul, S. and Furnas, G. W. (1997). Navigation in Electronic Worlds: A CHI 97 Workshop. SIGCHI Bulletin 29(4).

[KGE02]   Kaasten, S. and Greenberg, S. and Edwards, C. (2002) How People Recognize Previously Seen WWW Pages from Titles, URLs and Thumbnails. In X. Faulkner, J. Finlay, F. Detienne (Eds) People and Computers XVI (Proceedings of Human Computer Interaction 2002), BCS Conference Series, 247-265, Springer Verlag.

[KKP+01]   Kahan, J., Koivunen, M.-R., Prud'Hommeaux, E. and Swick, R. R. (2001). Annotea: An Open RDF Infrastructure for Shared Web Annotations. In Proceedings of the Tenth International World Wide Web Conference on World Wide Web. (pp. 623-632). Hong Kong. ACM Press.

[KC02]   Kelly, D. and Cool, C. (2002). The effects of topic familiarity on information search behavior. In Proceedings of the Second ACM/IEEE Joint Conference on Digital Libraries (JCDL '02), Portland, OR, 74-75.

[Kim01]   Kim, K. 2001. Implications of User Characteristics in Information Seeking on the World Wide Web. International Journal of Human-Computer Interaction, 13(3):323–340.

[Koe02]   Koehler, W. (2002). Web Page Change and Persistence -- A Four-Year Longitudinal Study. Journal of the American Society for Information Science and Technology, 53(2):162-171.

[Kuh91]   Kuhlthau, C. (1991). Inside the Search Process: Information Seeking from the User's Perspective. Journal of the American Society for Information Science 42(5):361-371.

[Law00]   Lawrence, S. (2000). Context in Web Search. IEEE Data Engineering Bulletin 23(3): 25-32.

[LBW00]   Lazonder, A., Biemans, H., and Wopereis, I. (2000). Differences between Novice and Experienced Users in Searching Information on the World Wide Web. Journal of the American Society for Information Science, 51(6):576-581.

[LNW99]   Lieberman, H., Nardi, B. A. and Wright, D. (1999). Training Agents to Recognize Text by Example. In Proceedings of the Third Annual Conference on Autonomous Agents. (pp. 116-122). 1999. ACM Press.

[MS98]   McDonald, S., and Stevenson, R.J. 1998. Effects of text structure and prior knowledge of the learner on navigation in hypertext. Human Factors, 40(1):18-27.

[MB97a]   Maglio, P. and Barrett, R. (1997). How to Build Modeling Agents to Support Web Searchers. In Proceedings of the Sixth International Conference on User Modeling (UM97, June 2-5, 1997). Springer Wien New York, 5-16.

[MB00]   Maglio, P. and Barrett, R. (2000). Intermediaries Personalize Information Streams. Communications of the ACM. 43: 96-101.

[MB97b]   Marais, H. and Bharat, K. (1997). Supporting Cooperative and Personal Surfing with a Desktop Assistant. In 10th Annual Symposium on User Interface Software and Technology. (pp. 129-138). ACM.

[Mal83]   Malone, T. W. 1983. How do people organize their desks?: Implications for the design of office information systems. ACM Trans. Inf. Syst. 1, 1 (Jan. 1983), 99-112.

[MDM+88]   Mayes, J. T., Draper, S. W., McGregor, A. M. and Oatley, K. (1988). Information Flow in a User Interface: the Effect of Experience and Context on the Recall of MacWrite Screens. In People and Computers IV. Cambridge University Press.

[MPC01]   Morrison, J., Pirolli, P., and Card, S. (2001). A Taxonomic Analysis of What World Wide Web Activities Significantly Impact People's Decisions and Actions. In Proceedings of the 2001 Conference on Human Factors in Computing Systems (CHI '01) Extended Abstracts. ACM Press, New York, 163-164.

[Mos98]   Mosaic User's Guide. (1998). Mosaic User's Guide: Annotations. Retrieved April 4, 2002, from <http://archive.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/help-on-annotate-win.html>.

[MB98]   Miller, J. R. and Bonura, T. (1998). From Documents to Objects. ACM SIGCHI Bulletin. 30: 53-58.

[NCS93]   NCSA Mosaic Documentation. (1993). NCSA Mosaic Documentation: Group Annotations in NCSA Mosaic (Mosaic 1.2). Retrieved April 4, 2002, from <http://archive.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/group-annotations.html>.

[NMW98]   Nardi, B. A., Miller, J. R. and Wright, D. J. (1998). Collaborative, Programmable Intelligent Agents. Communications of the ACM. 41: 96-104.

[NAM97]   Nestorov, S., Abiteboul, S. and Motwani, R. (1997). Inferring Structure in Semistructured Data. SIGMOD Record 26(4): 39-43.

[Nie06]   Nielsen, J. (2006). Search Engines and Leeches on the Web. Jakob Nielsen's Alertbox, January 9, 2006. Accessed on January 30, 2006 from: http://www.useit.com/alertbox/search_engines.html

[OC03]   Olston, C. and Chi, E. (2003). ScentTrails: Integrating Browsing and Searching on the Web. ACM Transactions on Computer-Human Interaction 10(3):177-197.

[PK97]   Pandit, M. S. and Kalbag, S. (1997). The Selection Recognition Agent: Instant Access to Relevant Information and Operations. In Proceedings of the 1997 International Conference on Intelligent User Interfaces. (pp. 47-52). Orlando, Florida, USA. ACM Press.

[Pay91]   Payne, S. J. (1991). Display-Based Action at the User Interface. International Journal of Man-Machines Studies 35: 275-289.

[Pew05a]   Pew Internet and American Life Project. 2005. Daily Internet Activities. Web page appears to have been created on December 05, 2005. Accessed on January 29, 2006 from: http://www.pewinternet.org/trends/Daily_Activities_12.05.05.htm

[Pew05b]   Pew Internet and American Life Project. 2005. Internet Activities. Web page appears to have been created on December 05, 2005. Accessed on January 31, 2006 from: http://www.pewinternet.org/trends/Internet_Activities_12.05.05.htm

[Pog04]   Pogue, D. (2004). Google Mail: Virtue Lies in the In-Box. New York Times, May 13, 2004. Accessed on March 2, 2006 from http://www.nytimes.com/2004/05/13/technology/circuits/13stat.html

[RS05]   Rainie, L., and Shermak, J. 2005. Pew Internet and American Life Project. Data memo on Search Engine use November 2005. Accessed on January 29, 2006 from: http://www.pewinternet.org/pdfs/PIP_SearchData_1105.pdf

[Ram05]   Ramakrishnan, N. (2005). The Traits of the Personable. In (B. Mobasher and S.S. Anand, Eds.) LNCS/LNAI State-of-the-Art Survey on Intelligent Techniques in Web Personalization, Vol. 3169 (LNAI Series), Springer-Verlag, pages 53-68.

[RSK04]   Ravasio, P., Schär, S. G., Krueger, H. (2004). In pursuit of desktop evolution: User problems and practices with modern desktop systems. ACM Transactions on Computer-Human Interaction, 11(2): 156-180.

[Rho00]   Rhodes, B. J. (2000). Margin Notes: Building a Contextually Aware Associative Memory. In Proceedings of the 2000 International Conference on Intelligent User Interfaces. (pp. 219-224). January 9-12, 2000. ACM.

[RS96]   Rhodes, B. J. and Starner, T. (1996). The Remembrance Agent: A continuously running automated information retrieval system. In Proceedings of The First International Conference on The Practical Application of Intelligent Agents and Multi Agent Technology (PAAM '96). (pp. 487-495). April 1996.

[Riz04]   Rizzo, T. (2004). "WinFS 101: Introducing the New Windows File System". Microsoft Corporation, March 17, 2004. Accessed on March 2, 2006 from http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnwinfs/html/winfs03112004.asp

[RCL+98]   Robertson, G., Czerwinski, M., Larson, K., Robbins, D. C., Thiel, D., and van Dantzich, M. (1998). Data mountain: using spatial memory for document management. In Proceedings of the 11th Annual ACM Symposium on User interface Software and Technology (San Francisco, California, United States, November 01 - 04, 1998). UIST '98. ACM Press, New York, NY, 153-162.

[SR03]   Shiri, A. and Revie, C. (2003). The effects of topic complexity and familiarity on cognitive and physical moves in a thesaurus-enhanced search environment. Journal of Information Science, 29(6):517-526.

[Spi03]   Spinellis, D. (2003). The decay and failures of Web references. Communications of the ACM, 46(1):71-77.

[Suc87]   Suchman, L. (1987). Plans and Situated Actions. Cambridge; New York: Cambridge University Press.

[TG97]   Tauscher, L. and Greenberg, S. (1997). How people revisit Web pages: empirical findings and implications for the design of history systems. International Journal of Human-Computer Studies 47(1):97-137.

[TAA+04]   Teevan, J., Alvarado, C., Ackerman, M. S. and Karger, D. (2004). The Perfect Search Engine Is Not Enough: A Study of Orienteering Behavior in Directed Search. In Proceedings of the 2004 Conference on Human Factors in Computing Systems. ACM Press, New York, 415-422.

[TFK+03]   Toms, E., Freund, L., Kopak, R., and Bartlett, J. (2003). The Effect of Task Domain on Search. In Proceedings of the 2003 Conference of the Center for Advanced Studies on Collaborative Research. IBM Press, 303-312. Availabile electronically through the ACM Guide to Computing Literature at http://portal.acm.org/

[Vak99]   Vakkari, P. (1999). Task complexity, problem structure and information actions: Integrating studies on information seeking and retrieval. *Information Processing and Management 35*(6):819-837.

[Wen03]   Wen, J. (2003). Post-Valued Recall Web Pages: User Disorientation Hits the Big Time. IT & Society, 1(3):184-194.

[WS96]   Whittaker, S. and Sidner, C. 1996. Email overload: exploring personal information management of email. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground (Vancouver, British Columbia, Canada, April 13 - 18, 1996). M. J. Tauber, Ed. CHI '96. ACM Press, New York, NY, 276-283.

[WH01]   Whittaker, S. and Hirschberg, J. (2001). The Character, Value, and Management of Personal Paper Archives. ACM Transactions Computer-Human Interaction 8(2): 150-170.

[WHN98]   Whittaker, S., Hirschberg, J., and Nakatani, C. H. (1998). Play it again: a study of the factors underlying speech browsing behavior. In CHI 98 Conference Summary on Human Factors in Computing Systems (Los Angeles, California, United States, April 18 - 23, 1998). CHI '98. ACM Press, New York, NY, 247-248.

[Wil04]   Wildemuth, B. 2004. The Effects of Domain Knowledge on Search Tactic Formulation. Journal of the American Society for Information Science and Technology, 55(3):246-258.

[WBL+06]   Wildemuth, B., Barry, C., Luo, L. Crystal, A., Oh, S. 2006. Establishing a Research Agenda for Studies of Online Search Behaviors: A Delphi Study. Accessed on January 27, 2006 from: http://ils.unc.edu/sig_use_delphi/

[Wil99]   Wilson, T.D. (1999). Models in Information Behaviour Research. Journal of Documentation, 55(3):249-270.

[WOR+71]   Witkin, M.A., Oltman, P., Raskin, E., and Karp, S. (1971, 2002). Group Embedded Figures Test Manual. Redwood City, CA: Mind Garden, Inc.

[WRM+02]   Woodruff, A., Rosenholtz, R., Morrison, J., Faulring, A., and Pirolli, P. (2002). A Comparison of the Use of Text Summaries, Plain Thumbnails, and Enhanced Thumbnails for Web Search Tasks. Journal of the American Society for Information Science and Technology, 53(2):172-185.

[W3C02]   World Wide Web Consortium. Semantic Web Development. Retrieved May 13, 2002, from <http://www.w3.org/2000/01/sw/>.

[Yee98]   Yee, K.-P. (1998). CritLink: Better Hyperlinks for the WWW. Retrieved May 13, 2002, from <http://crit.org/~ping/ht98.html>.

[Yee02a]   Yee, K.-P. The CritLink Mediator. Retrieved May 13, 2002, from <http://crit.org/critlink.html>.

[Yee02b]   Yee, K.-P. CritSuite. Retrieved May 13, 2002, from <http://crit.org>.

[ZG04]   Zhang, J. and Ghorbani, A. (2004). Familiarity and Trust: Measuring Familiarity with a Web Site. In Proceedings of the Second Annual Conference on Privacy, Security, and Trust (PST'04), October 13-15, 2004, Wu Centre, University of New Brunswick Fredericton, New Brunswick, Canada, pp. 23-28. Accessed on February 24, 2006 from http://dev.hil.unb.ca/Texts/PST/pdf/zhang.pdf

This page intentionally left blank.

Appendix A

Additional ANOVA Analyses

Many analysis techniques such as ANOVA rely on an assumption that the data being analyzed follows a normal distribution. In the data collected for this study, some of the dependent variables had skewed distributions, often to a ceiling or floor. In order to investigate this, each of the major analyses presented in this chapter was run once with the data as collected, and again with a log transformation applied to the data. Applying a log transform is a common technique used to adjust skewed data into a more normal distribution. A comparison of the two analyses – the original data and the log-transformed data – showed that almost all the statistical tests and effects were the same between the two analyses. A description of the log transform used and comparison of the log transform analyses to the analyses of the original data is presented in this Appendix.

The log transform was performed in SAS using the following code:

```
data tallstacked6;
set SASUSER.ALLSTACKED6;
        LATT = log(ATT);
        LNU  = log(NU);
        LNS  = log(NS+0.001);
        LSR  = log(SR+0.001);
        LKH  = log(KH);
        LHO  = log(HO);
        LAD  = log(AD);
        LBD  = log(BD);
run;
```

where:

- ATT is task completion time
- NU is number of URLs
- NS is number of searches
- SR is search ratio
- KH is task familiarity
- HO is task frequency
- AD is after-task difficulty rating
- BD is before-task difficulty rating

Note that since NS and SR could be zero, a small factor (0.001) was added to each before performing the log transform to avoid taking the log of zero.

A.1   Analysis 1 – Controlled Factors Analysis

See Section 5.2 for details of the analysis on the original data.

Table 45. Results of Two-Way ANOVA for Session x Task (EP-15 S12 tasks)

	Session (D)		Task (T)		Interaction (DxT)	
	F	*P*	*F*	*P*	*F*	*P*
Task Completion Time	20.06	<.01*	31.42	<.01*	2.61	<.01*
Number of URLs	9.10	0.01*	12.46	<.01*	1.78	0.06
Number of Searches	3.47	0.08	8.41	<.01*	0.64	0.80
Search Ratio	0.12	0.73	13.55	<.01*	0.20	1.00
Task Familiarity	49.80	<.01*	15.23	<.01*	3.60	<.01*
Task Frequency~	0.12	0.74	18.48	<.01*	0.53	0.88
Before-task Difficulty	37.88	<.01*	15.97	<.01*	3.29	<.01*
After-task Difficulty	21.98	<.01*	10.99	<.01*	1.23	0.28

* p ≤ 0.01, ~ see note about task frequency in section 4.4.4.1

Table 46. LOG Results of Two-Way ANOVA for Session x Task (EP-15 S12 tasks)

	Session (D)		Task (T)		Interaction (DxT)	
	F	*P*	*F*	*P*	*F*	*P*
Task Completion Time	15.29	<.01*	32.08	<.01*	1.47	0.15
Number of URLs	5.82	0.03**	16.10	<.01*	1.47	0.15
Number of Searches	1.22	0.29	19.47	<.01*	0.54	0.87
Search Ratio	0.54	0.47	20.35	<.01*	0.43	0.94
Task Familiarity	49.44	<.01*	12.22	<.01*	4.61	<.01*
Task Frequency~	0.03	0.87	22.24	<.01*	0.79	0.65
Before-task Difficulty	32.79	<.01*	13.33	<.01*	2.86	<.01*
After-task Difficulty	20.71	<.01*	9.99	<.01*	1.34	0.21

* p ≤ 0.01, ~ see note about task frequency in section 4.4.4.1

A.2   Analysis 2 – Prior Frequency Analysis

See Section 5.3 for details of the analysis on the original data.

Table 47. Results of Two-Way ANOVA for Session x Prior Frequency (EP-15 S12 tasks)

	Session (D)		Prior Freq (Q)		Interaction (DxQ)	
	F	P	F	P	F	P
Task Completion Time	12.27	<.01*	20.16	<.01*	3.60	0.04**
Number of URLs	7.11	0.02**	9.26	<.01*	2.16	0.13
Number of Searches	3.40	0.08	9.84	<.01*	0.36	0.70
Search Ratio	0.18	0.68	9.43	<.01*	0.39	0.68
Task Familiarity	43.76	<.01*	43.07	<.01*	8.29	<.01*
Before-task Difficulty	35.67	<.01*	45.46	<.01*	3.30	0.05**
After-task Difficulty	22.32	<.01*	20.48	<.01*	2.71	0.08

* p ≤ 0.01, ** p ≤ 0.05

Table 48. LOG Results of Two-Way ANOVA for Session x Prior Frequency (EP-15 S12 tasks)

	Session (D)		Prior Freq (Q)		Interaction (DxQ)	
	F	P	F	P	F	P
Task Completion Time	11.67	<.01*	20.62	<.01*	4.08	0.03**
Number of URLs	5.53	0.03**	18.27	<.01*	2.26	0.12
Number of Searches	1.31	0.27	20.55	<.01*	0.70	0.51
Search Ratio	0.66	0.43	18.89	<.01*	0.68	0.51
Task Familiarity	42.37	<.01*	32.95	<.01*	8.92	<.01*
Before-task Difficulty	30.87	<.01*	40.17	<.01*	3.38	0.05**
After-task Difficulty	20.28	<.01*	17.31	<.01*	2.11	0.14

* p ≤ 0.01, ** p ≤ 0.05

A.3   Analysis 3 – Prior Familiarity Analysis

See Section 5.4 for details of the analysis on the original data.

Table 49. Results of Two-Way ANOVA for Session x Prior Familiarity (EP-15 S12 tasks)

	Session (D)		Prior Fam (K)		Interaction (DxK)	
	F	P	F	P	F	P
Task Completion Time	13.29	<.01*	9.06	<.01*	4.25	0.02**
Number of URLs	7.56	0.01*	4.81	0.02**	5.18	0.01*
Number of Searches	4.80	0.04**	24.43	<.01*	1.71	0.20
Search Ratio	1.13	0.30	50.18	<.01*	1.28	0.29
Task Frequency~	0.59	0.45	65.37	<.01*	0.97	0.39
Before-task Difficulty	81.14	<.01*	141.39	<.01*	23.21	<.01*
After-task Difficulty	20.50	<.01*	26.58	<.01*	9.09	<.01*

* p ≤ 0.01, ** p ≤ 0.05, ~ see note about task frequency in section 4.4.4.1

Table 50. LOG Results of Two-Way ANOVA for Session x Prior Familiarity (EP-15 S12 tasks)

	Session (D)		Prior Fam (K)		Interaction (DxK)	
	F	P	F	P	F	P
Task Completion Time	11.71	<.01*	10.19	<.01*	4.71	0.02**
Number of URLs	6.35	0.02**	10.08	<.01*	5.97	0.01*
Number of Searches	5.33	0.03**	67.13	<.01*	3.29	0.05**
Search Ratio	3.65	0.07	70.07	<.01*	2.96	0.07
Task Frequency~	0.16	0.70	55.38	<.01*	0.31	0.74
Before-task Difficulty	76.98	<.01*	126.87	<.01*	23.68	<.01*
After-task Difficulty	19.41	<.01*	23.03	<.01*	8.32	<.01*

* p ≤ 0.01, ** p ≤ 0.05, ~ see note about task frequency in section 4.4.4.1

A.4   Analysis 4 – Task Similarity Analysis

See Section 5.5 for details of the analysis on the original data.

Table 51. Results of Two-Way ANOVA for Session x Task Similarity (EP-15 S12 tasks)

	Session (D)		Similarity (I)		Interaction (DxI)	
	F	*P*	*F*	*P*	*F*	*P*
Task Completion Time	12.27	<.01*	8.08	0.01*	1.64	0.22
Number of URLs	7.15	0.02**	0.40	0.54	0.89	0.36
Number of Searches	2.24	0.15	11.68	<.01*	0.27	0.61
Search Ratio	0.09	0.76	35.94	<.01*	0.46	0.51
Task Familiarity	28.21	<.01*	7.80	0.01*	2.64	0.12
Task Frequency~	0.04	0.84	0.79	0.39	0.40	0.54
Before-task Difficulty	22.82	<.01*	5.13	0.04**	1.25	0.28
After-task Difficulty	17.05	<.01*	1.34	0.26	0.29	0.59

* p ≤ 0.01, ** p ≤ 0.05, ~ see note about task frequency in section 4.4.4.1

Table 52. LOG Results of Two-Way ANOVA for Session x Task Similarity (EP-15 S12 tasks)

	Session (D)		Similarity (I)		Interaction (DxI)	
	F	*P*	*F*	*P*	*F*	*P*
Task Completion Time	8.34	0.01*	2.30	0.15	0.03	0.86
Number of URLs	4.69	0.05**	0.52	0.48	0.38	0.55
Number of Searches	1.43	0.25	31.19	<.01*	0.09	0.77
Search Ratio	0.74	0.40	36.74	<.01*	0.18	0.68
Task Familiarity	27.37	<.01*	8.86	0.01*	4.46	0.05**
Task Frequency~	0.00	0.98	2.79	0.11	0.74	0.40
Before-task Difficulty	19.88	<.01*	6.60	0.02**	1.85	0.19
After-task Difficulty	16.43	<.01*	1.27	0.28	0.50	0.49

* p ≤ 0.01, ** p ≤ 0.05, ~ see note about task frequency in section 4.4.4.1

A.5   Analysis 4 – Task Similarity Analysis of Session 2 tasks only

See Section 5.5.2 for details of the analysis on the original data.

Table 53. Results of One-Way ANOVA for Task Similarity (EPS-15 S2 tasks)

	Num DF	Den DF	Similarity (I) F	P
Task completion time	2	32	4.15	0.03**
Number of URLs	2	32	0.61	0.55
Number of searches	2	32	3.67	0.04**
Search ratio	2	32	6.97	<.01*
Task familiarity	2	32	12.31	<.01*
Before-task difficulty	2	32	10.50	<.01*
After-task difficulty	2	32	9.11	<.01*

* p ≤ 0.01, ** p ≤ 0.05

Table 54. LOG Results of One-Way ANOVA for Task Similarity (EPS-15 S2 tasks)

	Num DF	Den DF	Similarity (I) F	P
Task completion time	2	32	5.32	0.01*
Number of URLs	2	32	2.70	0.08
Number of searches	2	32	9.36	<.01*
Search ratio	2	32	9.62	<.01*
Task familiarity	2	32	10.85	<.01*
Before-task difficulty	2	32	9.37	<.01*
After-task difficulty	2	32	9.24	<.01*

* p ≤ 0.01, ** p ≤ 0.05

Appendix B

SAS Input and Output

In this appendix, sections of the SAS input and output files are presented. The entire set of input and output files is too long to include here, so only representative samples are included.

B.1   Analysis 1 – Controlled Factors Analysis, Selected Input

```
* =========================================;
* Means and Standard Errors;
* =========================================;


proc means data=SASUSER.ALLSTACKED6 mean std stderr t probt;
title 'CFA Means class=D  var=ATT NU NS SR KH HO AD BD w/EP-15 both D12';
class D;
var ATT NU NS SR KH HO AD BD;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;

proc means data=SASUSER.ALLSTACKED6 mean std stderr t probt;
title 'CFA Means class=FT  var=ATT NU NS SR KH HO AD BD w/EP-15 both D12';
class FT;
var ATT NU NS SR KH HO AD BD;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;

proc means data=SASUSER.ALLSTACKED6 mean std stderr t probt;
title 'CFA Means class=D FT  var=ATT NU NS SR KH HO AD BD w/EP-15 both D12';
class D FT;
var ATT NU NS SR KH HO AD BD;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;


* =========================================;
* Controlled factors analyses;
* =========================================;


*CFA-MIX-ATT;
proc mixed data=SASUSER.ALLSTACKED6;
title 'CFA-MIX-ATT: ALL6 DV=ATT == IV=D FT D*FT w/EP-15 both D12';
options formdlim='-';
class P D FT;
model ATT=D FT D*FT;
lsmeans D FT D*FT / pdiff=all adjust=tukey;
repeated / subject=P type=cs;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;

*CFA-MIX-NU;
proc mixed data=SASUSER.ALLSTACKED6;
title 'CFA-MIX-NU: ALL6 DV=NU == IV=D FT D*FT w/EP-15 both D12';
options formdlim='-';
class P D FT;
```

```
model NU=D FT D*FT;
lsmeans D FT D*FT / pdiff=all adjust=tukey;
repeated / subject=P type=cs;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;

*CFA-MIX-NS;
proc mixed data=SASUSER.ALLSTACKED6;
title 'CFA-MIX-NS: ALL6 DV=NS == IV=D FT D*FT w/EP-15 both D12';
options formdlim='-';
class P D FT;
model NS=D FT D*FT;
lsmeans D FT D*FT / pdiff=all adjust=tukey;
repeated / subject=P type=cs;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;

*CFA-MIX-SR;
proc mixed data=SASUSER.ALLSTACKED6;
title 'CFA-MIX-SR: ALL6 DV=SR == IV=D FT D*FT w/EP-15 both D12';
options formdlim='-';
class P D FT;
model SR=D FT D*FT;
lsmeans D FT D*FT / pdiff=all adjust=tukey;
repeated / subject=P type=cs;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;

*CFA-MIX-KH;
proc mixed data=SASUSER.ALLSTACKED6;
title 'CFA-MIX-KH: ALL6 DV=KH == IV=D FT D*FT w/EP-15 both D12';
options formdlim='-';
class P D FT;
model KH=D FT D*FT;
lsmeans D FT D*FT / pdiff=all adjust=tukey;
repeated / subject=P type=cs;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;

*CFA-MIX-HO;
proc mixed data=SASUSER.ALLSTACKED6;
title 'CFA-MIX-HO: ALL6 DV=HO == IV=D FT D*FT w/EP-15 both D12';
options formdlim='-';
class P D FT;
model HO=D FT D*FT;
lsmeans D FT D*FT / pdiff=all adjust=tukey;
repeated / subject=P type=cs;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;

*CFA-MIX-AD;
proc mixed data=SASUSER.ALLSTACKED6;
title 'CFA-MIX-AD: ALL6 DV=AD == IV=D FT D*FT w/EP-15 both D12';
options formdlim='-';
class P D FT;
model AD=D FT D*FT;
lsmeans D FT D*FT / pdiff=all adjust=tukey;
repeated / subject=P type=cs;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;

*CFA-MIX-BD;
proc mixed data=SASUSER.ALLSTACKED6;
title 'CFA-MIX-BD: ALL6 DV=BD == IV=D FT D*FT w/EP-15 both D12';
options formdlim='-';
class P D FT;
model BD=D FT D*FT;
lsmeans D FT D*FT / pdiff=all adjust=tukey;
repeated / subject=P type=cs;
where NUSE EQ 1 AND ( I EQ "E" OR I EQ "P" ) AND FT NE 15;
run;
```

B.2   Analysis 1 – Controlled Factors Analysis, Selected Output

```
-------------------------------------------------------------------------------------------
                    CFA-MIX-ATT: ALL6 DV=ATT == IV=D FT D*FT w/EP-15 both D12              160
                                                               16:08 Tuesday, October 18, 2005

                                The Mixed Procedure

                                 Model Information

            Data Set                      SASUSER.ALLSTACKED6
            Dependent Variable            ATT
            Covariance Structure          Compound Symmetry
            Subject Effect                P
            Estimation Method             REML
            Residual Variance Method      Profile
            Fixed Effects SE Method       Model-Based
            Degrees of Freedom Method     Between-Within


                             Class Level Information

            Class     Levels    Values

            P            17     12 13 14 15 16 17 18 20 21 22
                                23 25 26 29 30 31 32
            D             2     1 2
            FT           12     1 3 5 6 8 10 11 12 13 16 17 18


                                    Dimensions

                    Covariance Parameters           2
                    Columns in X                   39
                    Columns in Z                    0
                    Subjects                       17
                    Max Obs Per Subject            24


                             Number of Observations

                Number of Observations Read          345
                Number of Observations Used          345
                Number of Observations Not Used        0


                                Iteration History

        Iteration    Evaluations    -2 Res Log Like        Criterion

               0            1        3511.07982089
               1            2        3488.04987710        0.00000001
               2            1        3488.04985950        0.00000000


                             Convergence criteria met.
-------------------------------------------------------------------------------------------
                    CFA-MIX-ATT: ALL6 DV=ATT == IV=D FT D*FT w/EP-15 both D12              161
                                                               16:08 Tuesday, October 18, 2005

                                The Mixed Procedure

                          Covariance Parameter Estimates

                          Cov Parm     Subject     Estimate

                          CS           P            361.16
                          Residual                 2344.82


                                  Fit Statistics

                    -2 Res Log Likelihood          3488.0
                    AIC (smaller is better)        3492.0
```

```
         AICC (smaller is better)        3492.1
         BIC (smaller is better)         3493.7


              Null Model Likelihood Ratio Test

            DF     Chi-Square      Pr > ChiSq

             1        23.03           <.0001


            Type 3 Tests of Fixed Effects

                 Num      Den
       Effect     DF       DF     F Value    Pr > F

       D            1       16      20.06     0.0004
       FT          11      151      31.42     <.0001
       D*FT        11      138       2.61     0.0048
```

Appendix C

Preliminary Survey

This appendix contains screen capture images of the on-line preliminary survey described in Chapter 3.
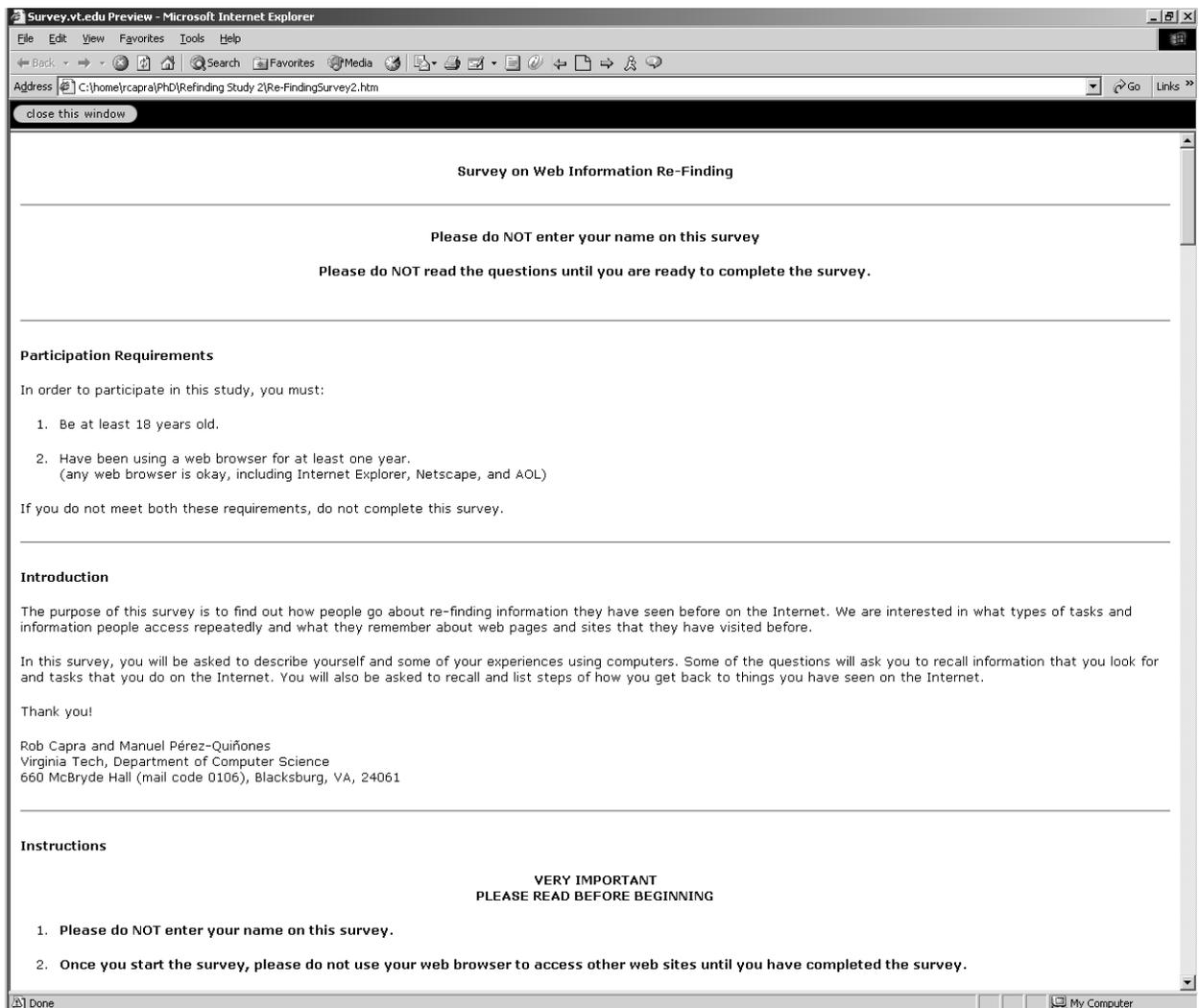
Browser window — Survey.vt.edu Preview – Microsoft Internet Explorer

Address: C:\home\rcapra\PhD\Refinding Study 2\Re-FindingSurvey2.htm

close this window

3. Please do NOT read any questions until you are ready to complete the survey.

4. The survey has three sections that should be completed in order.

5. Complete the survey yourself. Do NOT get help from other people.

Important Information

Your participation is completely voluntary and you are free not to answer any questions that you do not wish to respond to. If you do choose to complete this survey, you will help our research on how to design tools to help people re-find information on the web.

Your responses will be anonymous to the extent that you should NOT write your name or other information that would uniquely identify you on the survey. Numbers, not names, will be used to index the data collected from the surveys. The results of this study and the information you provide on the survey may be published, including direct quotations of your responses. In reports of the results, no participant names will be included.

Agreement to Participate

I have read and understand the Instructions and Important Information above and meet the requirements to participate in this study. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent.

☐ **Check here to indicate your agreement**

Section 1

1. Age

2. Gender

     ○ Male   ○ Female

3. Occupation

     ○ Undergraduate student
     ○ Graduate student
     ○ University faculty
     ○ University staff

17. Things you do INFREQUENTLY (e.g. less often than daily or weekly) using a web browser:

Task or Information Looked For / Approximate Frequency (check one)

A. [_____]

    ○ A few times a month   ○ Once per month   ○ Less than once per month

B. [_____]

    ○ A few times a month   ○ Once per month   ○ Less than once per month

C. [_____]

    ○ A few times a month   ○ Once per month   ○ Less than once per month

D. [_____]

    ○ A few times a month   ○ Once per month   ○ Less than once per month

E. [_____]

    ○ A few times a month   ○ Once per month   ○ Less than once per month

---

Section 3 -- Web Actions

IMPORTANT -- Please read all the instructions on this page.

In this section, we would like you to list the steps you take to perform some of the items you listed in Section 2. In particular, we will ask you to do this for the first three items from each question in Section 2.

You will need to look back at Section 2 to see what each of these items was, but please do not change any of your responses in Section 2.

If some of your tasks from Section 2 are general in nature such as "on-line shopping", you may wish to use a specific instance of the task when you write down your list of steps in this section (i.e. think of the steps you take to use a particular on-line shopping site).

Example #1
If one of your responses in Section 2 was "check to see if the world record

Section 3 -- Web Actions

IMPORTANT -- Please read all the instructions on this page.

In this section, we would like you to list the steps you take to perform some of the items you listed in Section 2. In particular, we will ask you to do this for the <u>first three items</u> from each question in Section 2.

You will need to look back at Section 2 to see what each of these items was, but please do not change any of your responses in Section 2.

If some of your tasks from Section 2 are general in nature such as "on-line shopping", you may wish to use a specific instance of the task when you write down your list of steps in this section (i.e. think of the steps you take to use a particular on-line shopping site).

> <u>Example #1</u>
> If one of your responses in Section 2 was "check to see if the world record
> for pogo stick jumping has changed," you might list the following steps for it
> in this section:
>
>> a. Go to www.guinnessworldrecords.com by typing in the URL
>> b. Type in "pogo stick" in the search box on the left of the Guinness page
>> c. Look at the results for the pogo stick record

Please try to be as specific as possible. If you can't remember something exactly, indicate the partial information that you do remember, or list the steps you would take to get to the result based on what you do remember.

> <u>Example #2</u> From the example above, if you remembered that the web site was for
> the Guinness Book of World Records, but could not remember the exact URL,
> you might list the following steps:
>
>> a. Search for "Guinness Book" on a search engine (please indicate which one)
>> b. Look at the results for a link to the Guinness World Records web site
>> c. Click on that link
>> d. Type in pogo stick in the search box on the left of the Guinness page
>> e. Look at the results for the pogo stick record

Please type in how you would do the task using your own web browser (e.g. if you have shortcuts or bookmarks that you would use, please indicate this).

If you don't remember anything about the steps you take, that is okay, just type "Don't remember" on the first line and go on to the next item.

18. Please list the steps for Question #16, Item A:

a.
b.
c.
d.

Appendix D

Visual Basic Data Collection Program

This appendix contains screen capture images of the custom Visual Basic data collection program used to present the task descriptions and to collect information from participants about each task. The program is described in section 4.2. Note that the task description shown in these images is a sample task that was not used in the actual study.

Instructions

In the tasks for today, you will be asked to look for some information using the Internet and a web browser.

You will also be asked to answer some questions before and after each task.

For some of the tasks, you may be asked to think out loud to describe what you are thinking and doing as you work on the task.

There are a total of 18 tasks for today's session. You may work at the tasks for up to 1 hour total.

Please work through the tasks in order.

Next

Instructions (continued)

In the tasks for today, you will be asked to look for some information using the Internet and a web browser.

If you cannot find the information asked for by a task, that is okay. You can stop work on the task and go to the next one.

Go about the tasks as you normally would use the web to find information.

If you normally create or use bookmarks, short-cuts, or other types of notes, please feel free to do so.

When you are ready to begin, please click the button below to start the first task.

Start

Before Task Questions   _ □ ✕

Task 1

Read the task below, and **before you start working on it,** answer the following questions.

Task Description

    Go to a web page where you can login to read your email
    (if you know of one)

---

1. How often do you do this type of task using the Internet in your day-to-day life?

○	○	○	○	○	○	○	○
More than once per day	Once per day	Several times per week	Once per week	Several times per month	Once per month	Less often than once per month	Never

2. How difficult do you think this task will be?

○	○	○	○	○	○	○
Very difficult	Difficult	Somewhat difficult	Average	Somewhat easy	Easy	Very easy

3. How well do you know how to get to the information?

○	○	○	○	○	○	○
1	2	3	4	5	6	7

I Don't know
  where it is

I Know exactly
  where it is

4. Is this task focused on finding a specific piece of information, or on browsing around to find something of interest?

○	○	○	○
1	2	3	4

Finding specific
information

Browsing for
something of interest

5. Speaking out loud, briefly describe the strategy you plan to use to get to the information.

[Next]

After Task Questions

Task 1

6. How likely are you to need to get back to this type of information in the future?

○	○	○	○	○	○	○
Very likely	Likely	Somewhat likely	Average	Somewhat unlikely	Unlikely	Very unlikely

7. How important is this type of information to you?

○	○	○	○	○	○	○
Very important	Important	Somewhat important	Average	Somewhat unimportant	Unimportant	Very unimportant

8. How difficult was this task?

○	○	○	○	○	○	○
Very difficult	Difficult	Somewhat difficult	Average	Somewhat easy	Easy	Very easy

9. Speaking out loud, describe why the task was easy or difficult?

10. Speaking out loud, describe what (if anything) would have helped you find the information faster or more easily?

Next

After Task Questions

Task 1

6. How difficult was this task?

○ Very difficult   ○ Difficult   ○ Somewhat difficult   ○ Average   ○ Somewhat easy   ○ Easy   ○ Very easy

7. **Speaking out loud,** describe why the task was easy or difficult?

8. **Speaking out loud,** describe what helped you find the information.

9. **Speaking out loud,** describe what you remembered from last time about finding this or similar information?

10. **Speaking out loud,** describe what you remembered that you didn't use or that was confusing?

11. Was it easier to get the information this time?

○ Much easier   ○ Easier   ○ Somewhat easier   ○ About the same   ○ Somewhat harder   ○ Harder   ○ Much harder

12. **Speaking out loud,** describe what (if anything) would have helped you find the information faster or more easily?

Finish

Appendix E

Finding and Refinding Tasks

This appendix contains tables of the finding and refinding tasks, with the full task descriptions used in the study and a list of what information was requested for each task.

This page intentionally left blank.

Table 55. Session 1 and Session 2 Tasks Compared

Task Num	Topic	Session 1 Task Description	Similarity	Session 2 Task Description
1	Dictionary	Look up the meaning of the word, salubrious	Path	Look up the meaning of the word, frugivorous
2	ReadNews	Find and READ news story about world news	M	Find the news story that you read last week about world news.
3	Phone	Find the phone number for the Mish Mish art supply store in downtown Blacksburg	Exact	Find the phone number for the Mish Mish art supply store in downtown Blacksburg
4	Shopping	Look for prices on Women's Sweatshirts. Find two possibilities that you might buy for a gift to a friend or relative. Please look at at least three different sweatshirts from three different places to buy.	Subset	Find one of the Women's Sweatshirts that you looked at during the session last week.
5	Travel	Find a flight from Roanoke, VA on Dec. 1 to Chicago, IL and return on Dec. 4	Path	Find a flight from Roanoke, VA on Dec. 14 to Chicago, IL and return on Dec. 19
6	Research	Find instructions about how to solve the Rubik's cube.	Exact	Find instructions about how to solve the Rubik's cube.
7	Travel	Imagine that you are planning a trip to San Francisco, California. Find two tourist attractions or events that you might be interested in. Please look at at least four different atrractions or events.	Subset	Find one of the tourist attractions or events for San Francisco, California that you looked at during the session last week.
8	Phone	Find the phone number for the Kroger grocery store in Blacksburg near the Math Emporium	Path	Find the phone number for the Kroger grocery store in Blacksburg on South Main Street near El Guadalupe
9	Homepage	Go to the Wal-Mart Web site	Different	Go to the Web site for Dr. David Patterson at the University of California at Berkeley
10	Sports	Find the score from the most recent Atlanta Falcon's NFL football game.	Exact	Find the score from the most recent Atlanta Falcon's NFL football game.
11	Movies	Find a movie that you are interested in seeing. Find a list of the starting times for the movie for today.	Exact	Find a movie that you are interested in seeing. Find a list of the starting times for the movie for today.
12	Stocks	Find the current price for Microsoft stock.	Exact	Find the current price for Microsoft stock.
13	Homepage	Go to the Web site of Dr. Stasko at Georgia Tech	Exact	Go to the Web site of Dr. Stasko at Georgia Tech
14	Weather	Find tomorrow's hi and low temperature for Madrid, Spain	Different	Go to the Web site for The Record Exchange music store on Main Street in Blacksburg
15	Research	Find an instruction manual for a Panasonic AG-1340 Video Cassette Recorder (VCR)	Exact	Find an instruction manual for a Panasonic AG-1340 Video Cassette Recorder (VCR)
16	Weather	Find tomorrow's hi and lo temperature for Blacksburg	Exact	Find tomorrow's hi and lo temperature for Blacksburg
17	Academic	Find out what the last day of classes is for the Spring 2004 semester	Path	Find out when Spring Break is for the Spring 2004 semester
18	Map	Find out how to get from the Blacksburg Municipal Golf Course to the Macados on University Blvd.	Exact	Find out how to get from the Blacksburg Municipal Golf Course to the Macados on University Blvd.

Table 56. Session 1 Task Descriptions and Information Requested

Task Num	Session 1 Task Description	Information Format	Information Requested
1	Look up the meaning of the word, salubrious	Text box	Brief definition
2	Find and READ news story about world news	Radio buttons Yes, no	Did you find and read one?
3	Find the phone number for the Mish Mish art supply store in downtown Blacksburg	Text box	Phone number
4	Look for prices on Women's Sweatshirts. Find two possibilities that you might buy for a gift to a friend or relative. Please look at at least three different sweatshirts from three different places to buy.	Text box Text box Text box Text box	Sweatshirt #1, name and brand Sweatshirt #1, Why did you choose this one? Sweatshirt #2, name and brand Sweatshirt #2, Why did you choose it?
5	Find a flight from Roanoke, VA on Dec. 1 to Chicago, IL and return on Dec. 4	Text box Text box Text box Text box	Airline name Flight times Ticket price Why did you choose this one?
6	Find instructions about how to solve the Rubik's cube.	Radio buttons Yes, no	Did you find it?
7	Imagine that you are planning a trip to San Francisco, California. Find two tourist attractions or events that you might be interested in. Please look at at least four different atrractions or events.	Text box Text box Text box Text box	Event #1, name Event #1, Why did you choose this one? Event #2, name Event #2, Why did you choose this one?
8	Find the phone number for the Kroger grocery store in Blacksburg near the Math Emporium	Text box	Phone number
9	Go to the Wal-Mart Web site	Radio buttons Yes, no	Did you find it?
10	Find the score from the most recent Atlanta Falcon's NFL football game.	Text box Text box	Opponent name Game score
11	Find a movie that you are interested in seeing. Find a list of the starting times for the movie for today.	Text box Text box	Movie title Earliest movie time
12	Find the current price for Microsoft stock.	Text box	Stock price
13	Go to the Web site of Dr. Stasko at Georgia Tech	Radio buttons Yes, no	Did you find it?
14	Find tomorrow's hi and low temperature for Madrid, Spain	Text box Text box	Hi Lo
15	Find an instruction manual for a Panasonic AG-1340 Video Cassette Recorder (VCR)	Radio buttons Yes, no	Did you find it?
16	Find tomorrow's hi and lo temperature for Blacksburg	Text box Text box	Hi Lo
17	Find out what the last day of classes is for the Spring 2004 semester	Text box	Date
18	Find out how to get from the Blacksburg Municipal Golf Course to the Macados on University Blvd.	Text box	Estimated distance

Table 57. Session 2 Task Descriptions and Information Requested

Task Num	Session 2 Task Description	Information Format	Information Requested
1	Look up the meaning of the word, frugivorous	Text box	Brief definition
2	Find the news story that you read last week about world news.	Radio buttons Yes, no	Did you find it?
3	Find the phone number for the Mish Mish art supply store in downtown Blacksburg	Text box	Phone number
4	Find one of the Women's Sweatshirts that you looked at during the session last week.	Text box Text box	Sweatshirt, name and brand Why did you go back to this one?
5	Find a flight from Roanoke, VA on Dec. 14 to Chicago, IL and return on Dec. 19	Text box Text box Text box Text box	Airline name Flight times Ticket price Why did you choose this one?
6	Find instructions about how to solve the Rubik's cube.	Radio buttons Yes, no	Did you find it?
7	Find one of the tourist attractions or events for San Francisco, California that you looked at during the session last week.	Text box Text box	Event name Event price/cost
8	Find the phone number for the Kroger grocery store in Blacksburg on South Main Street near El Guadalupe	Text	Phone number
9	Go to the Web site for Dr. David Patterson at the University of California at Berkeley	Radio buttons Yes, no	Did you find it?
10	Find the score from the most recent Atlanta Falcon's NFL football game.	Text box Text box	Opponent name Game score
11	Find a movie that you are interested in seeing. Find a list of the starting times for the movie for today.	Text box Text box	Movie title Earliest movie time
12	Find the current price for Microsoft stock.	Text box	Stock price
13	Go to the Web site of Dr. Stasko at Georgia Tech	Radio buttons Yes, no	Did you find it?
14	Go to the Web site for The Record Exchange music store on Main Street in Blacksburg	Radio buttons Yes, no	Did you find it?
15	Find an instruction manual for a Panasonic AG-1340 Video Cassette Recorder (VCR)	Radio buttons Yes, no	Did you find it?
16	Find tomorrow's hi and lo temperature for Blacksburg	Text box Text box	Hi Lo
17	Find out when Spring Break is for the Spring 2004 semester	Text box Text box	Start date End date
18	Find out how to get from the Blacksburg Municipal Golf Course to the Macados on University Blvd.	Text box	Estimated distance

This page intentionally left blank.

Appendix F

Institutional Review Board Approval Memos

This appendix contains copies of the Virginia Tech Institutional Review Board Approval Memos for the studies conducted as part of this dissertation.

03-371

Virginia Tech
VIRGINIA POLYTECHNIC INSTITUTE
AND STATE UNIVERSITY

Institutional Review Board

Dr. David M. Moore
IRB (Human Subjects) Chair
Assistant Vice Provost for Research Compliance
CVM Phase II - Duckpond Dr., Blacksburg, VA 24061-0442
Office: 540/231-4991; FAX: 540/231-6033
e-mail: moored@vt.edu

August 11, 2003

MEMORANDUM

TO: Manuel Pérez-Quiñones Computer Science 0106
             Robert Capra Computer Science 0105

FROM: David M. Moore

SUBJECT: IRB EXEMPTION APPROVAL – " Survey on Web Information Re-
             Finding" – IRB #03-371

I have reviewed your request to the IRB for exemption for the above referenced project.
I concur that the research falls within the exempt status.   Approval is granted effective as
of August 7, 2003.

cc: file

Virginia
Tech
VIRGINIA POLYTECHNIC INSTITUTE
AND STATE UNIVERSITY

Institutional Review Board

Dr. David M. Moore
IRB Chairman Subject Chair
Assistant Vice Provost for Research Compliance
CVM Phase II - Duckpond Dr., Blacksburg, VA 24061-0442
Office: 540/231-4991; FAX: 540/231-0035
e-mail: moored@vt.edu

November 12, 2003

MEMORANDUM

TO: Manuel A. Perez-Quinones Computer Science 0106
             Robert Capra Computer Science 0106

FROM: David M. Moore

SUBJECT: **Expedited Approval – "An Investigatin of Web Information Re-
             Finding"– IRB #03-539**

This memo is regarding the above-mentioned protocol. The proposed research is eligible
for expedited review according to the specifications authorized by 45 CFR 46.110 and 21
CFR 56.110. As Chair of the Virginia Tech Institutional Review Board, I have granted
approval to the study for a period of 12 months, effective November 11, 2003.

Approval of your research by the IRB provides the appropriate review as required by
federal and state laws regarding human subject research. It is your responsibility to
report to the IRB any adverse reactions that can be attributed to this study.

To continue the project past the 12 month approval period, a continuing review
application must be submitted (30) days prior to the anniversary of the original approval
date and a summary of the project to date must be provided. My office will send you a
reminder of this (60) days prior to the anniversary date.

Cc: File

*A Land-Grant University—The Commonwealth Is Our Campus
An Equal Opportunity / Affirmative Action Institution*