2012

Exploring the Effect of Cognitive Load on the Propensity for Query Reformulation Behavior

Kyoungsik Na

THE FLORIDA STATE UNIVERSITY

COLLEGE OF COMMUNICATION AND INFORMATION

EXPLORING THE EFFECT OF COGNITIVE LOAD ON THE PROPENSITY FOR QUERY REFORMULATION BEHAVIOR

By

KYOUNGSIK NA

A Dissertation submitted to the
School of Library and Information Studies
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Summer Semester, 2012

Kyoungsik Na defended this dissertation on April 03, 2012.

The members of the supervisory committee were:

Kathleen Burnett
Professor Directing Dissertation

Neil Charness
University Representative

Gary Burnett
Committee Member

Besiki Stvilia
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.

To my parents

Seungwoo Na and Injae Park

I dedicated this dissertation to my father and mother with respect, appreciation, and love and to the God without whom none of my achievements would be possible.

ACKNOWLEDGEMENTS

I would like to express my sincere and deep gratitude to Dr. Kathleen Burnett who has been served as my major professor, advisor, and mentor for her limitless advice, support, encouragement, and guidance during my entire doctoral program. I also would like to thank my committee member: Dr. Gary Burnett, Dr. Besiki Stvilia, and Dr. Neil Charness. Without them this dissertation would not be possible.

I would like to give my special thanks to Dr. Neil Charness for his advice, guidance, encouragement, and suggestions for developing my research and completing it.

I also would like to give my deepest appreciation to my parents, Seungwoo Na and Injae Park, and my entire family members for their endless supports and prayers throughout the years of my study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

With the aim of improving information retrieval system design, this study explored the effect of cognitive load on the propensity to reformulate queries during information seeking on the Web, specifically the effect of manipulating three affective components that contribute to cognitive load—mental demand, temporal demand, and frustration

A significant difference in the propensity of query reformulation behavior was found between searchers exposed to cognitive load manipulations and searchers who were not exposed. Those exposed to cognitive load manipulations, namely, *mental demand*, *temporal demand*, and *frustration*, made 2.18 times fewer search queries than searchers not exposed. Furthermore, the NASA-TLX cognitive load scores of searchers who were exposed to the three cognitive load manipulations were higher than those of searchers who were not exposed. However, the propensity of query reformulation behavior did not differ across task types. The findings suggest that a dual-task method and NASA-TLX assessment serve as good indicators of cognitive load. Because the findings show that cognitive load hinders a searcher's interaction with information search tools, this study concludes by recommending strategies for reducing cognitive load when designing information systems, or user interfaces.

CHAPTER ONE

INTRODUCTION

Background

The purpose of this study was to investigate the effect of cognitive load on the propensity to reformulate queries during information seeking on the World Wide Web (the Web), specifically the effect of manipulating three affective components of cognitive load—mental demand, temporal demand, and frustration. This research is informed by cognitive load theory (CLT).

In the past decade searching for information on the Web has increased dramatically. According to Pew Internet & American Life Project tracking surveys (2008), 89% of internet users report using search engines making search one of most popular internet activities. Pew Internet survey in May 2011 further finds that 92% of online adults use search engines to find information on the Web, including 59% who do so on a typical day (2011). The Web has become an indispensable tool for finding information in everyday life. However, due to information overload (Toffler, 1970) on the Web, searching for relevant information may not always be enjoyable. Finding appropriate and accurate information is often time-consuming and frustrating. Even though searchers can find information about most topics on the Web, they may not find what they need in a timely fashion; they may not know where to look or how to find what they need.

Information seeking with a search engine is a complex, evolving (Bates, 1989), dynamic, (White & Roth, 2009) interactive, and iterative process (Pirolli & Card, 1999; Rieh & Xie, 2006). Information seeking is a learning process requiring many physical and mental demands including: information processing, problem-solving, interpretation, comparison, adaptation, and judgment. Furthermore, various factors such as information search system features, the nature of search tasks, the time allocated for searching, and searcher characteristics may affect one's ability to find relevant information (Bell & Ruthven, 2004; Bystrom & Jarvelin, 1995; Vakkari, 1999).

In addition to searcher characteristics, the nature of the search tasks and the difficulty level may affect search strategies. Search tasks can be divided into two categories: open-ended questions and closed questions (Navarro-Prieto, Scaife, & Rogers, 1999; White & Iivonen, 2001). According to Wilkinson, Reader, and Payne (2012), task difficulty and time pressure affect information foraging strategies during adaptive browsing. Piroli (2007) says that information foraging strategies for well-structured problems differ from ill-structured problems. Piroli (2007) also states that the "information environment is a potential source of valuable knowledge that can improve our ability to achieve our goals, especially when they involve ill-structured tasks" (p.21).

In an attempt to find more relevant information on the Web, searchers can reformulate or expand queries (Belkin et al., 2001; Efthimiadis, 1996a, 1996b; Jansen, Zhang, & Spink, 2007; Rieh & Xie, 2001, 2006; White & Roth, 2009). Previous research has rigorously explored patterns, transitions, and transformations of query reformulation in order to explain how such actions were executed during a search cycle (Belkin, et al., 2001; Bruza, Dennis, & McArthur, 2000; Efthimiadis, 1996a, 1996b; Jansen, Zhang, et al., 2007; Rieh & Xie, 2001, 2006; Sihvinen & Vakkari, 2004; Silverstein, Henzinger, Marais, & Moricz, 1999; Spink, Jansen, & Ozmultu, 2001; Stojanovic, 2005; White & Roth, 2009). Query formulation and reformulation may increase the time spent on information seeking. Ideally, it would be unnecessary to engage in information searching if the appropriate information would always be at the information seeker's fingertips. However, it is rarely the case that information seekers initially receive the best possible search results. Therefore, it is necessary to approach the study of the information seeking process (ISP) and query reformulation as an iterative activity (Kuhlthau, 1991; Marchionini, 1995).

Jansen, Spink, and Saracevic (2000) find that web searchers use relatively short queries and review only a few pages of results. Hearst (2009) notes that when searchers choose a way to express information needs that do not match relevant documents, they become "reluctant to radically modify their original query" and get "stuck on the original formulation" (p.141). Some studies (Belkin, et al., 2001; Efthimiadis, 1996b; Jansen, Zhang, et al., 2007) find that query reformulation is a popular search tactic. Researchers have developed techniques to automatically generate thesauri to support query reformulation (Jansen, 2006b; Jansen & McNeese, 2005). Google, Bing, and Yahoo provide on-the-spot assistance to help searchers generate query formulation and reformulation. Nonetheless, it is reasonable to question why searchers are often reluctant to reformulate queries in information retrieval tasks even though studies (Hearst, 2009) show that query reformulation may draw better search results.

Web searching involves the use of working memory resources to process information as learning (Ingwersen, 1996, 2000). Human cognition is involved in information processing during the search process (Spector, Merrill, Merrienboer, & Driscoll, 2007; Vakkari, 2000; Vakkari, Pennanen, & Serola, 2003). However, to date, few studies have investigated how cognitive load affects searchers' experiences during information searching (Back & Oppenheim, 2001; Bruza, et al., 2000; Dennis, Bruza, & McArthur, 2002; Dennis, McArthur, & Bruza, 1998; Gwizdka, 2008; Rieh & Xie, 2001, 2006). Gwizdka (2010) states that cognitive load is significantly higher during query formulation and tagging of relevant documents as compared to examining search results and viewing individual documents. Gwizdka (2010) also notes that "understanding what contributes to a user's cognitive load during search tasks is crucial to understanding the search process and to identifying which search task types and search system features make greater demands of users" (p. 2167).

In educational and cognitive psychology, cognitive load is understood to play an important role in terms of facilitating learning and performance (Chandler & Sweller, 1991; Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Sweller, 1988, 1994; Sweller, Chandler, Tierney, & Cooper, 1990), but research on information retrieval in Library and Information Studies has not given this sufficient attention. The term, "cognitive load," or simply "load" has several synonymous terms such as "mental workload," "cognitive workload," or simply "workload."

There is still a lack of research specifically exploring the effect of cognitive activities—such as mental demand, temporal demand, effort, and frustration—on cognitive load, by examining the search processes of end searchers. This study will explore what happens to the propensity for searchers to reformulate queries when three components of cognitive load—mental demand, temporal demand, and frustration—are manipulated.

Research Background

The principles of CLT take into account the limitations of human working memory capacity (Miller, 1956). CLT has important implications for instructional design. According to CLT, if the resources of working memory are exceeded, learning will be ineffective (Sweller, et al., 1990). Although the effects of cognitive load have not been explored rigorously in information retrieval research, cognitive load has been the subject of extensive study in educational psychology and applied learning sciences research. In these studies, multimodality of learning methods has been shown to reduce cognitive load to achieve maximum learning.

Gwizdka (2010) examines the distribution of cognitive load in web searching. He finds that cognitive load is distributed based upon characteristics of the search task types and system features. Gwizdka (2010) also demonstrates that cognitive load is significantly higher during query formulation and user description of relevant documents as compared to during examination of search results and viewing individual documents. Xie and Salvendy (2000) explore predictors of mental workload in single and multiple task environments. Instantaneous workload, average workload, accumulated workload, peak workload and overall workload were examined. Kim and Rieh (2005) state that mental effort is dynamic and scattered throughout search tasks in library and web searching.

In this study, the researcher assumes that as a searcher's cognitive load increases, his or her ability to perform effective web searching slowly decreases until the searcher reaches the point of cognitive overload, at which point the searcher may choose to give up the task, present emotional instability, or lose motivation to pursue the goal.

Statement of the Problem

Previous information retrieval studies have largely explored the effectiveness and efficiency of search systems (Grossman & Frieder, 2005; Hu, Ma, & Chau, 1999; Jones, 1981; Marchionini, 1995; Raghavan & Jung, 1989; Voorhees & Harman, 2005) without paying much attention to the cognitive load of the searcher. Since searchers expend energy through cognitive activities during information retrieval (Ingwersen, 1992, 1996, 2000), understanding the cognitive aspects of information seeking is important to the development of efficient and effective search systems.

Based on the research in educational and cognitive psychology described above, this study assumes that as a searcher's cognitive load increases, his or her ability to perform effectively slowly decreases as it approaches the point of cognitive overload. Once searchers reach the point of cognitive overload, they may choose to give up on a particular search, present emotional instability, or lose motivation to pursue their original information seeking goal. This study addresses whether or not there is any difference in the propensity for query reformulation in variations of cognitive load between groups when each of the following three affective components of cognitive load are manipulated: mental demand, temporal demand, and frustration.

Significance of the Study

It is expected that the findings of this study will contribute to the development and implementation of information systems and services that will lead to more effective Web searching behavior. This study also seeks to understand query reformulation with three long-term goals:

1) To improve search engine design and performance to better meet searcher needs;
2) To better understand searchers' query reformulation behavior; and,
3) To improve query formulation support and education for searchers.

This study expects to promote understanding of the cognitive load of searchers in order to be of better assistance to both searchers and system developers.

Research Questions

The following research question and three sub-questions guide this study.

Research Question: Is there any difference in the propensity for query reformulation behavior between searchers who experience cognitive load manipulation and searchers who do not experience cognitive load manipulation?

Sub-question 1: Is there any difference in the propensity for query reformulation behavior when mental demand increases?

Sub-question 2: Is there any difference in the propensity for query reformulation behavior when temporal demand increases?

Sub-question 3: Is there any difference in the propensity for query reformulation behavior when frustration increases?

These questions are described in more detail in Chapter 3.

Limitations of the Study

There are four factors that limit the generalizability of this study's results: 1) search engine characteristics; 2) searcher characteristics (including searcher knowledge level characteristics); 3) lack of query reformulation; and 4) effectiveness of the information retrieval system.

Search Engine Characteristics

All search engines have unique characteristics. Some include keyword searching capability only, others combine keyword searching with command language search capabilities, and others employ natural language search capabilities. Therefore, the results from one type cannot necessarily be generalized to other types. Google is free and is the most often used search engine. It provides specific search features, including: phrase search (""), search within a specific website (site:), term exclusion (-), fill in the blanks (*), search exactly as stated (+), the OR operator, and several advanced features (e.g., weather for many US and worldwide cities "weather" followed by the name of the city or zip code; and definition of a word or phrase, "define" followed by the word or phrase).

Searcher Characteristics

The subjects in this study were students from the College of Communication and Information at Florida State University. Results of this study are not generalizable to other populations, including student populations in other disciplines. Novice and experienced searchers are assumed to have different information seeking behaviors (Hsieh-Yee, 1993); therefore, the results of this study may not be generalizable to novice searchers. Future research may look at the differences between novice and experienced searchers.

Lack of Query Reformulation

This research studied query reformulation. Previous research (Klink, 2001a) demonstrates that searchers often do not attempt to reformulate queries even when query reformulation is necessary to find relevant results. Bruza, Dennis, and McArthur (2000) claimed that short queries are a well-known activity on the Web search.

Effectiveness of Information Retrieval Systems

This study did not evaluate the effectiveness of information retrieval systems or search engines; rather, it focused on the searchers' information seeking behavior (i.e., query reformulation behaviors differed by imposed cognitive load).

Summary

This chapter introduced the problem, significance, research questions, and limitations of the proposed study. The next chapter reviews the previous relevant research literature.

CHAPTER TWO

LITERATURE REVIEW

Introduction

This chapter reviews the relevant literature on cognitive load, query reformulation, and relevance feedback. It begins with a brief introduction to search experience and the general problems of information seeking on the Web that have been previously discussed. Then, it presents detailed background on cognitive load theory. The next section explains the definition of “query” and “query reformulation.” The concluding section discusses relevance feedback.

Search Experience

With the growing use of the Web for information retrieval, interaction between system and user has grown in importance. A query is an information need representing the concept or the topic that the searcher wants to examine or know more about. A query can be transformed to another query that the searcher reformulates from the previous query and inserts into the system in order to look for more relevant information (Hearst, 2009). The ideal search would be one input with one best matching output for the user’s information need. However, searchers do not get information from just one source; they pick up bits of information from many different sources while they explore. Bates (1989) refers to this process as “berrypicking.” According to Bates (1989)

> at each stage [of the information search process], with each different conception of the query, the user may identify useful information and references … the query is satisfied not by a single final retrieved set, but by a series of selections of individual references and bits of information at each stage of the ever-modifying search (p. 410).

The search process begins with the information need. This information need is expressed by the searcher who will structure a concept and formulate it into a search strategy (i.e., a search query that can be understood by the information retrieval system). If a search is completed with an initial search query, there is no need to reformulate the previous query. Thus, the ideal search experience would include retrieval of the desired information with minimal effort at the time the search is performed, but this is not often achieved in real information seeking situations. According to Case (2002), “the principle of least effort, which is chiefly pragmatic and not at all optimal, predicts that seekers will minimize the effort required to obtain information, even if it means accepting a lower quality or quantity of information” (p. 143).

One of the searcher’s burdens in information retrieval is query formulation and reformulation. Searchers typically are not experts at formulating queries on information systems with not very intuitive interfaces and when the retrieval yields poor results. Manning, Raghavan, and Schutze (2008) describe the characteristics of the Web search users as follows.

> … web search users tend to not know (or care) about the heterogeneity of web content, the syntax of query languages and the art of phrasing queries; indeed, a mainstream tool (as web search has come to become) should not place such onerous demands on billions of people. A range of studies has concluded that the average number of keywords in a web search is somewhere between 2 and 3. Syntax operators (Boolean connectives, wildcards, etc.) are seldom used, again a result of the composition of the audience – “normal” people, not information scientists (p. 432).

Another burden for searchers is evaluating the initial search results and determining whether these results contain information that will best meet the information need, or if a subsequent query must be formulated to achieve this goal. It is not always possible for searchers to formulate initial queries that produce optimal or even acceptable search results, and thus they may need to reformulate queries until the desired information is located. Query reformulation is therefore often a necessary component of the information seeking process which might better be seen as a continuum rather than a solitary act. Query reformulation may occur for the reasons including the following.

1) The searcher does not find adequate and relevant information;
2) Retrieved search results are too few;
3) Retrieved search results are too many;
4) Retrieved search results are not satisfying to the searcher;

5) The searcher is not confident about the retrieved search results;
6) The searcher wants to compare between search results.

Information Seeking on the Web

People who use internet rigorously engage in searching for information using a search engine as one of their common activities on the Web (Haythornthwaite, 2001). Search engines such as Google, Yahoo, or Bing on the Web adopted a certain information retrieval algorithm that attempts to efficiently and effectively identify documents for users' search queries that are most relevant to their information need. The need to obtain revenue from user of their services is pushing these companies to offer both advertising and search results in a co-mingled fashion. However, due to the vast amount of information and questionable authority of information on the Web, locating the most relevant information there can present challenges to searchers. One big challenge might be difficulty in judging relevance judgment due to information overload. To find more relevant information, searchers might increase the amount of time spent searching and using cognitive activities for relevant judgment, but this increases demands on the user during problem solving, reasoning, or thinking possibly affecting her general satisfaction and performance when completing a search (Schmutz, Heinz, Metrailler, & Opwis, 2009).

According to Case (2008), information behavior "encompasses information seeking as well as the totality of other unintentional or passive behaviors (such as glimpsing or encountering information), as well as purposive behaviors that do not involve seeking, such as avoiding information" (p. 5). Marchionini and Shneiderman (1988) state that information seeking depends on several factors including the seeker, task, search system, domain, setting, and search outcomes. According to Case (2008), information seeking is "a conscious effort to acquire information in response to a need or a gap in your knowledge" (p. 5). Case (2002) also writes, "human information seeking simply is not so simple. It is neither straightforward nor typically complete; it is more like a series of interruptions, punctuated by other interruptions" (p. 328).

Information seeking on the Web is an iterative, continuous, recursive, and complex process (Bates, 1989, 1990; Rieh & Xie, 2006) that may involve query reformulations to better meet the information needs of searchers. Bruza, Dennis, and McArthur (2000) claim that "query reformulation using Hyperindex Browser (HiB) does significantly improve the relevance of the documents … but the improvement in document relevance comes at the cost of increased search time and increased cognitive load" (p. 286). Previous studies (Jansen, et al., 2000) indicate that searchers are often reluctant to view more than the first page or screen of results. Other research (Aula, 2003) discusses how experience using computers, the web, and web search engines affect the query formulation process, while domain expertise did not have an effect on the query formulation. Interestingly, Aula (2003) claims that experienced users formulated longer and more specific queries whereas less experienced users formulated fewer and more generic queries. Hearst (2009) states that "at times, when a searcher chooses a way to express an information need that does not successfully match relevant documents, the searcher becomes reluctant to radically modify their original query and stays stuck on the original formulation" (p.141).

Cognitive Load Theory

Cognitive load theory (CLT) assumes that human beings have a limited working memory capacity that is connected to unlimited long-term memory (Baddeley, 1992). Miller (1956) posits that our capacity for processing information in working memory is limited to "the magic number seven plus or minus two.[1]" Cowan (2001) argues that today the number of short-term memory is much smaller than seven plus or minus two, which is a short-term memory of about four. Also, most people recognize that effective working memory contains a component known as long-term working memory (Ericsson & Kintsch, 1995). As a result of this limitation of working memory, CLT is concerned with controlling working memory capacity by imposing adequate levels of cognitive load (Sweller, 1999).

Recently, cognitive load theory has become one of the fundamental theories used to understand and explain cognitive activities in the learning process with learning technologies, especially in multimedia environments and online learning (Mayer, 2001). Research on cognitive load has identified possible sources that contribute to cognitive load and has found ways of reducing different types of cognitive load that are helpful to design instructional materials and strategies (Chandler & Sweller, 1991; Mayer, 2001; Sweller, van Merrienboer, & Pass, 1998).

[1] Miller's theory argues that on average the typical human mind can hold 7 plus or minus 2 items in his working memory.

To this end, CLT has been increasingly used to better understand cognitive load and develop instructional design of learning environments (Brunken & Leutner, 2001).

According to Pass and Merrienboer (1994), cognitive load is the total load that cognitive activities impose on the working memory during the completion of a task or learning. Sweller (1988) is well known for developing CLT, which differentiates three types of cognitive load (Sweller, van Merrienboer & Paas, 1998): intrinsic cognitive load, extraneous cognitive load, and germane cognitive load. Intrinsic load is imposed by the learning task; extraneous load by the learning environment in relation to the learning task being executed; germane load by the learner as he or she attempts to understand the learning materials or reach the learning goal. Therefore, germane load is often considered a "good" form of load. During a task or learning situation, the instructor's ultimate goal is to reduce intrinsic and ECL while generating GCL (Pass, Tuovinen, Tabbers, & Van Gerven, 2003). It is considered that manipulating the nature of a task or learning situation is difficult, but rather we could try to decrease extrinsic cognitive load and increase germane cognitive load. Simply put, the total cognitive load is the sum of intrinsic cognitive load, extraneous cognitive load, and GCL(Sweller, 1988). Paas, Renkl, and Sweller (2003) state that "intrinsic, extraneous, and germane cognitive loads are additive in that, together, the total load cannot exceed the working memory resources available if learning is to occur" (p. 2).

Types of Cognitive Load

As stated earlier, CLT in educational settings differentiates three types of cognitive load -- intrinsic, extraneous, and germane cognitive load (Sweller, Van Merrienboer, & Pass, 1998).

1) Intrinsic cognitive load (ICL) often involves the difficulty of the learning task that cannot be easily altered. Managing ICL is preferred to facilitate learning and performance since it has a negative relationship with learning (Moreno & Park, 2010, p. 17). In the context of information seeking, if a search task itself is too complex to understand or execute, then ICL would be expected to be correspondingly high.
2) Extraneous cognitive load (ECL) may be caused by the format of the information presentation or inadequate instructional design. Reducing ECL is preferred to facilitate learning and performance, since it has a negative relationship with learning (Moreno & Park, 2010, p. 17). In the context of information seeking, if the presentation format of

search results works well for a searcher, then ECL would be expected to be correspondingly low.

3) Germane cognitive load (GCL) is often defined as the effort of the learner to understand the learning materials or reach the learning goal. Inducing GCL is preferred to gain maximum learning and performance, since it has a positive relationship with learning (Moreno & Park, 2010, p. 17). In the context of information seeking, if the effort of the searcher increases during a search task, then GCL would be expected to be correspondingly high.

Cipperfield (2006) states that ICL for a given problem or task cannot be changed. However ECL and GCL can be adjusted and they are inversely proportional to each other.

Related Studies of Cognitive Load Theory

Cognitive load theory has been widely used in instructional technology and educational settings to facilitate learning and performance by minimizing ECL, managing ICL, and thereby promoting GCL. Although the concept of cognitive load has been discussed primarily in educational and psychological settings where the reduction of cognitive load is preferable so that the best learning and performance can be achieved, it should also be considered in information retrieval research to increase understanding of cognitive aspects of search behavior that can improve searcher performance, satisfaction, usability of websites, and interactivity of interfaces.

Sweller (1988) states that, with respect to CLT, "optimum learning occurs when the load on working memory is kept to a minimum to best facilitate information transfer to long-term memory"(p. 157). Cognitive load theory has many implications for the design of learning materials, which can help to keep the cognitive load of learners to a minimum during the learning process (Sweller, 1988, 1994). Mayer (2001) later applied CLT to the field of multimedia learning in that reducing ECL is preferable to facilitate learning with multimedia.

Sweller, Van Merrienboer, and Paas (1998) point out that optimum learning and problem-solving happen when a learner's cognitive process is controlled by minimizing ECL and maximizing GCL. Chandler and Sweller (1991) stated that when intrinsic and extraneous load are higher, the response time will be lower and the number of errors will be greater. Xie and

Salvendy (2000) reveal that mental workload was significantly affected by time pressure. Mental workload is a load that is imposed on human beings to interact between a task and human capabilities or resources. Mental workload is a multifaceted and multidimensional phenomenon and can be related to many facets such as physiological states, mental effort, physical effort, time pressure, performance and more (Wickens, 1992). O'Donnell and Eggemeier (1986) define workload as the portion of the human operator's limited capacities or resources that are required to perform a particular task.

Galy, Cariou, and Melan (2011) describe the relationship between mental workload factors and cognitive load types as follows.

> … mental workload studies revealed that the sensitivity of workload measures differs according to a number of factors, and in particular according to the cognitive task to be performed. This led to the proposal that several different mental workload categories should be distinguished, as has been suggested by Sweller …. the author distinguished three categories of cognitive load (p. 2).

Past studies (Chandler & Sweller, 1991; Sweller, et al., 1998) show that direct investigation of cognitive load is difficult because there are multiple variables and complex relationships between performance, mental load, and mental effort. As a consequence, cognitive load can often be estimated indirectly using alternative variables such as total estimated time, retention, error rate, outcome, and performance.

In examining user interface designs of information retrieval systems, Hu, Ma, and Chau (1999) use cognitive load as a measurement of the information seeking and processing effort and examined how searchers facilitate information gain by reducing cognitive load to increase searchers' satisfaction. Back and Oppenheim's (2001) study found that users would prefer an interface design requiring a relatively low cognitive load that at the same time, can result in high user satisfaction.. In their study, a self-reporting method was used to obtain user assessments of the cognitive load associated with a particular interface.

Assessment of Cognitive Load

When exploring cognitive load, researchers face the difficulty of assessing invisible cognitive load. According to Brunken, Plass, and Leutner (2003), cognitive load can be treated "as a theoretical construct, describing the internal processes of information processing that cannot be observed directly" (p.55). They (Brunken, et al., 2003) demonstrate the methods that could be used as a measurement of cognitive load in two dimensions: *objectivity* and *causal relation* (See Table 2.1).

Table 2.1

Two Dimensions of Measurement of Cognitive Load Assessment

	Causal Relationship	
Objectivity	Indirect	Direct
Subjective	Self-reported invested mental effort	Self-reported stress level Self-reported difficulty of materials
Objective	Physiological measures Behavioral measures Learning outcome measure	Brain activity measure Dual-task performance

Note. Adapted from "Direct Measurement of Cognitive Load in Multimedia Learning," by R. Brunken, J. L. Plass, and D. Leutner, 2003, *Educational Psychologist*, 38, p. 55.

Different methods have been used to measure cognitive load for different purposes. Questionnaires, think-aloud protocols, and interviews are considered indirect ways to assess cognitive load while methods like eye-tracking, haptics, and dual-task techniques that are using heavily external devices are more direct ways to assess cognitive load. Kim and Rieh (2005) employ a dual-task method to measure cognitive load in searching a library system and the Web. They found that the evaluation of search results consumed a different level of cognitive load from that of query formulation that impedes performance of a secondary task in a dual-task method. Gwizdka (2010) points out that "a variation of the dual-task methodology is used to show how cognitive load is sensitive to the dynamic changes in task demands such as the changes of load from one stage to another" (p. 2167). Gwizdka (2010) demonstrates that cognitive load is significantly higher during query formulation and users' description of a relevant document than other stages of the information search process as compared to during examination of search results and viewing individual documents.

CLT from both the psychological perspective and educational settings comprises the theoretical framework in this study. Cognitive load in psychology can be defined as workload imposed on a user while interacting with a system or task. However, the basic principle of this theory is worth mentioning because it is closely related to the performance of searchers.

Query and Query Reformulation

Users tend to come up with irrelevant search results without providing a proper query to a an information retrieval system (Jansen, Zhang, et al., 2007). Marchionini and White (2007) state that the quality of queries has a big effect on the quality of the search results. Therefore, queries are crucial to the successful search results. Spink, Wolfram, Jansen, & Saracevic (2001) (A Spink, Wolfram, Jansen, & Saracevic, 2001) reveal that most people use few search terms, few modified queries, view few Web pages, and rarely use advanced search features. Manning, Raghavan, and Schutze (2008, p. 432-433) describedthree types of web search queries:

1) Informational queries seek general information on a broad topic such as "Jaguar" or "Car." There is typically no a single web page that contains all the information sought; indeed, users with informational queries typically try to assimilate information from multiple web pages.
2) Navigational queries seek the website or home page of a single entity that the user has in mind, such as Lufthansa. In such cases, the user's expectation is that the very first search result should be the home page of Lufthansa. The user is not interested in a plethora of documents containing the term Lufthansa; for such a user, the best measure of user satisfaction is precision rather than recall. Precision is the fraction of a search output that is relevant for a particular query. The recall on the other hand is the ability of a retrieval system to obtain all or most of the relevant documents in the collection. In other words, precision is the number of relevant documents a search retrieves divided by the total number of documents retrieved, while recall is the number of relevant documents retrieved divided by the total number of existing relevant documents that should have

been retrieved (Shafi & Rather, 2005).

3) Transactional queries reflect the intent of the user performing a transaction on the Web, such as purchasing an item, downloading a file or making a reservation. In such cases, the search engine should return results listing services that provide form interfaces for such transactions.

According to Baeza-Yates and Ribeiro-Neto (1999), "a query is the formulation of a user information need. In its simplest form, a query is composed of keywords and the documents containing such keywords are searched for. Keyword-based queries are popular because they are intuitive, easy to express, and allow for fast ranking." (p. 100)(p. 100) Keyword-based searches are common in commercial search engines such as Google, Yahoo!, and Bing.

Query formulation (QF) is the first interaction taken when beginning a search using a search engine. Query reformulation (QR) is an iterative process that follows up on the QF and may continue until the end of the search process. There are ways to strategize query formulation and reformulation. For example, searchers may use Boolean operators, fields, Natural Language, proximity, thesauri, query by example, query by instances, or relevance feedback. The efficacy of each strategy is dependent upon two factors: 1) the features of the information search system, and 2) the individual's information seeking behavior characteristics.

Past research studied and analyzed user queries and QR behavior on the Web (Jansen, et al., 2000). Past studies also explored the effectiveness and patterns of user queries using search log analysis (Jansen, 2006a; Jansen & Spink, 2005a, 2005b; Jansen, Spink, Blakely, & Koshman, 2006; Jansen, Spink, & Koshman, 2007; Jansen, Spink, & Pedersen, 2005) and transitions between queries (Belkin, et al., 2001; Jansen, Zhang, et al., 2007; Klink, 2001a; Rieh & Xie, 2001, 2006; A. Spink, et al., 2001). Jansen, Zhang, and Spink (2007) found that about half of initial queries were modified during the search process and most of these were reformulated to be more specific queries. They categorized query reformulation patterns such as "new," "reformulation," "specialization," "content change," and "generalization." Jansen and McNeese (2005; 2006b) found that users will accept automated assistance to improve the quality of the search process, especially after viewing results and locating relevant documents.

According to Bruza and Dennis (1997), a query reformulation is the substitution, addition, or deletion of terms from a previous query, while continuing to include some of the terms from the previous query. Klink (2001b)states that searchers may reformulate queries many times until they obtain better results in cases where the results were insufficient or did not produce anything related to the searcher's needs.

Jansen, Zhang, and Spink (2007) point out that query patterns are modified by subtracting a noun or adding a verb after the initial query term during query reformulation. They also found that when modifying queries, searchers are more likely to add or remove keywords from previous queries by virtue of system assistance in which about half of the initial queries were modified and most were refined to be more specific.

Zazo, Figuerola, Berrocal, and Rodriguez (2005) measure the influence of the number of terms in the initial query. Their findings show that that when a user submits a large number of terms in the initial query, the possibility for improving results on the next query is low since long queries are more likely to match the description of a user's information need better than that of short queries; therefore, the longer queries leave less room for improvement.

White and Marchionini (2007) test the effectiveness of real-time query expansion (RTQE) in regard to task completion time, satisfaction with a search system, quality of results, and quality of queries. The results show that RTQE yields better quality initial queries, facilitates more engagement in the search, and increases query expansion. Other researchers have implemented the following techniques to help searchers reformulate queries more successfully (Hearst, 2009):

1) Spelling suggestions and corrections;
2) Automated term suggestions;
3) Suggesting popular destinations; and
4) Relevance feedback.

Relevance Feedback

The relevance of search results is important because it helps searchers assess whether or not they need to continue searching for the information they need. The quality of a query usually determines the level of relevance of search results. Searchers much engage their own cognition to decide whether search results are relevant. Gwizdka (2010) found that cognitive load is different between search task stages in that "query formulation and tagging of the relevant results were associated with higher cognitive load than viewing results lists and viewing individual content documents" (p. 2181).

Manning, Raghavan, & Schutze (2008) state that "a document is relevant if it addresses the stated information need, not because it just happens to contain all the words in the query. This distinction is often misunderstood in practice, because the information need is not overt. But, nevertheless, an information need is present" (2008, p. 152). Saracevic (1996) distinguishes five dimensions of relevance:

1) System (algorithm) relevance;
2) Topical (subject) relevance;
3) Cognitive (pertinence) relevance;
4) Situational (utility) relevance; and
5) Motivational (affective) relevance.

One interesting idea behind Saracevic's conceptualization of relevance is that these five dimensions can be categorized into two major aspects: system-oriented relevance and searcher-oriented relevance. Saracevic (2007a, 2007b) also points out that searchers may judge the relevance of information by assessing the topicality, aboutness, utility, novelty of, and satisfaction with that information.

Schamber, Eisenberg, and Nilan (1990) discuss the characteristics of relevance-as-concept:

1) Relevance is a multidimensional cognitive concept;
2) Relevance is a dynamic concept;
3) Relevance is a complex but systematic and measurable phenomenon.

Relevance feedback has been used attempts to effectively improve retrieval results (Salton & Buckley, 1990). Relevance feedback is a popular information retrieval tool (Grossman & Frieder, 2005) that takes advantage of searcher relevance judgments in the retrieval process. The basic idea behind relevance feedback is that a searcher inputs a query and the system returns a set of results so that the searcher can determine whether or not the results are relevant; the searcher subsequently reformulates the query or ends the search. Figure 2.1 illustrates multiple processes in relevance feedback.

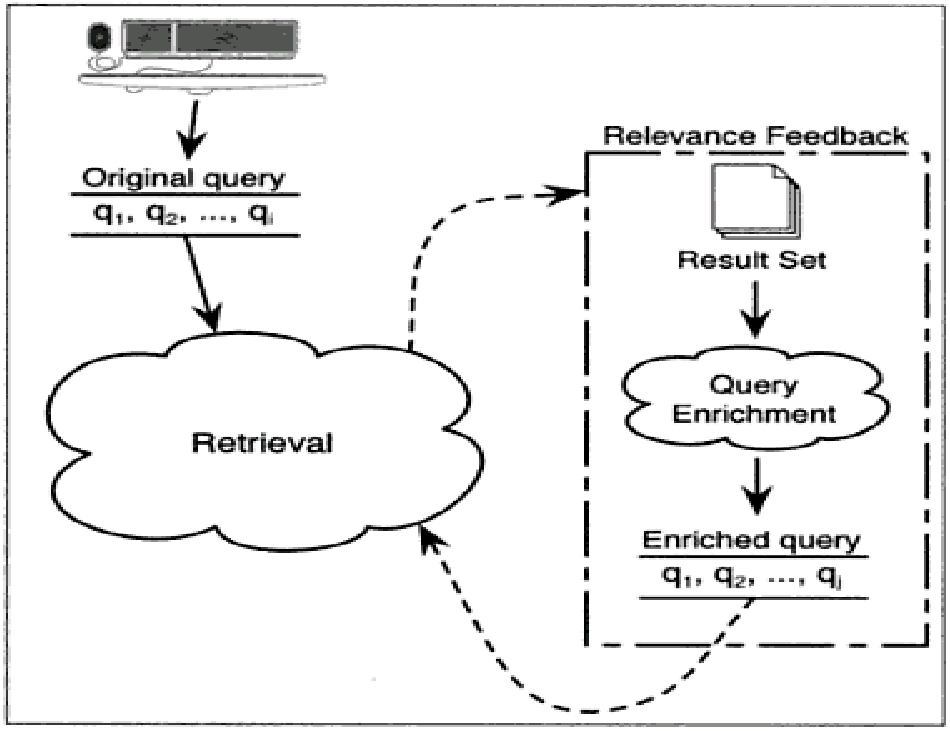

Figure 2.1. Relevance Feedback Process. Adapted from "Information retrieval: algorithms and heuristics," by D. A. Grossman and O.Frieder, 2005, p. 95.

Grossman and Frieder (2005) state that

> the basic premise is to implement retrieval in multiple passes. The user refines the query in each pass based on results of previous queries. Typically, the user indicates which of the documents presented in response to an initial query are relevant, and new terms are added to the query based on this selection. Additionally, existing terms in the query can be re-weighted based on user feedback (p.94).

Summary

This chapter reviewed the previous relevant research literature. The next chapter introduces the methodology used in this study.

CHAPTER THREE

METHODOLOGY

Introduction

The purpose of this study is to explore whether or not there is any difference in the propensity for query reformulation behavior when cognitive load is manipulated. This study employs an experimental design with pre-and post-surveys and search log analysis to analyze the effect of manipulations of cognitive load on the propensity for query reformulation between an experimental and a control group. Key components of the methodology include the following.

1) A pre-task questionnaire providing assessments of subjects to gain an understanding of demographics, computer knowledge, web search experience, and prior experience with Google.
2) An experiment allowing manipulations on three components that contribute to cognitive load imposed on subjects who perform a search task.
3) A post-task questionnaire providing assessments of NASA-TLX workload to gain assessment of cognitive load after a search task and assessment of overall query reformulation experience.
4) A search transaction log providing assessments of the propensity of query reformulations.

Research Questions

This study asks the following research question and sub-questions.

Research Question: Is there any difference in the propensity for query reformulation behavior between searchers who experience cognitive load manipulation and searchers who do not experience cognitive load manipulation?

Sub-question 1: Is there any difference in the propensity for query reformulation behavior when mental demand increases?

Sub-question 2: Is there any difference in the propensity for query reformulation behavior when temporal demand increases?

Sub-question 3: Is there any difference in the propensity for query reformulation behavior when frustration increases?

Hypotheses

This study tests the following hypotheses.

1) There are significant differences in the propensity for query reformulation behavior between searchers who experience cognitive load manipulation and searchers who do not experience cognitive load manipulations.
2) There is a significant difference in the propensity for query reformulation behavior when mental demand increases relative to those without mental demand.
3) There is a significant difference in the propensity for query reformulation behavior when temporal demand increases relative to those without temporal demand.
4) There is a significant difference in the propensity for query reformulation behavior when frustration increases relative to those without frustration.

Research Procedure

This research consists of four phases:

I. Phase 1: Experiment
II. Phase 2: Data analysis
III. Phase 3: Data coordination and data interpretation
IV. Phase 4: Conclusions and implications.

The experiment was administered in Phase 1, which included a pre-task survey, three search tasks, and a post-task survey at the time of each task. In Phase 2, the data from the pre-and post-task surveys and recorded search tasks were analyzed. Different data analysis tools were employed for each phase of the analysis. The researcher coordinated and interpreted the data in Phase 3. In Phase 4, the implications of the results were discussed and conclusions were drawn.

Experiment Procedure

The experiment consisted of the following activities.

1. Introduction to the search engine and a practice session;
2. Break;
3. Completion of the pre-task questionnaire;
4. Completion of search task 1;
5. Completion of a post-task questionnaire;
6. Break;
7. Completion of search task 2;
8. Completion of a post-task questionnaire;
9. Break;
10. Completion of search task 3;
11. Completion of a post-task questionnaire.

First, the experimenter briefly introduced the purpose of the study, the Google search engine, and provides a practice session. After the practice session, a pre-task survey was administered after which the search tasks began. Subjects were asked to complete the three search tasks. Following the completion of each search task, subjects were asked to complete a post-task questionnaire. Between search tasks, subjects were given a break to avoid carryover effects.

Experimental Design

The experimental design was chosen in light of the importance of comparing an experimental group to a control group when attempting to identify differences in the propensity for query reformulation behavior in past Information Retrieval research (Voorhees & Harman, 2005). The data collection methods employed in this study included a pre-survey (see Appendix A), experiment, a post-survey (see Appendix B), and search transaction logs.

Experiments are extremely helpful in assessing cause and effect relationships and can allow for the deliberate manipulation of one variable while keeping other variables constant (Creswell, 2003; Creswell & Plano-Clark, 2007; Goodwin, 2009).

Table 3.1

Groups and Manipulations

Group	Number of Subjects	Manipulation
Experimental Group	27	Mental Demand, Temporal Demand, and Frustration
Control Group	27	No manipulations
Total Subjects	54	

Experiment Conditions

To examine whether or not there is any difference in the propensity for query reformulation behavior among subjects, the researcher exposed each of the two groups of subjects to different stimuli. As shown in Table 3.1, a total of 54 subjects were recruited from the School of Library and Information Studies at Florida State University. The experimental group received three manipulations increasing cognitive load. The control group received no manipulation. Subjects in the control group were expected to experience relatively low to average cognitive load throughout the information search tasks, while subjects in the experimental group were expected to experience alteration in their cognitive loads during the search tasks due to the manipulations.

The researcher deliberately manipulated three components of cognitive load (mental demand, temporal demand, and frustration) to observe how specific variations in the levels of these three components affect the propensity for query reformulation behavior between the two groups. Each participant was asked to subjectively rate the cognitive load at the end of each task situation (See Figure 3.1).

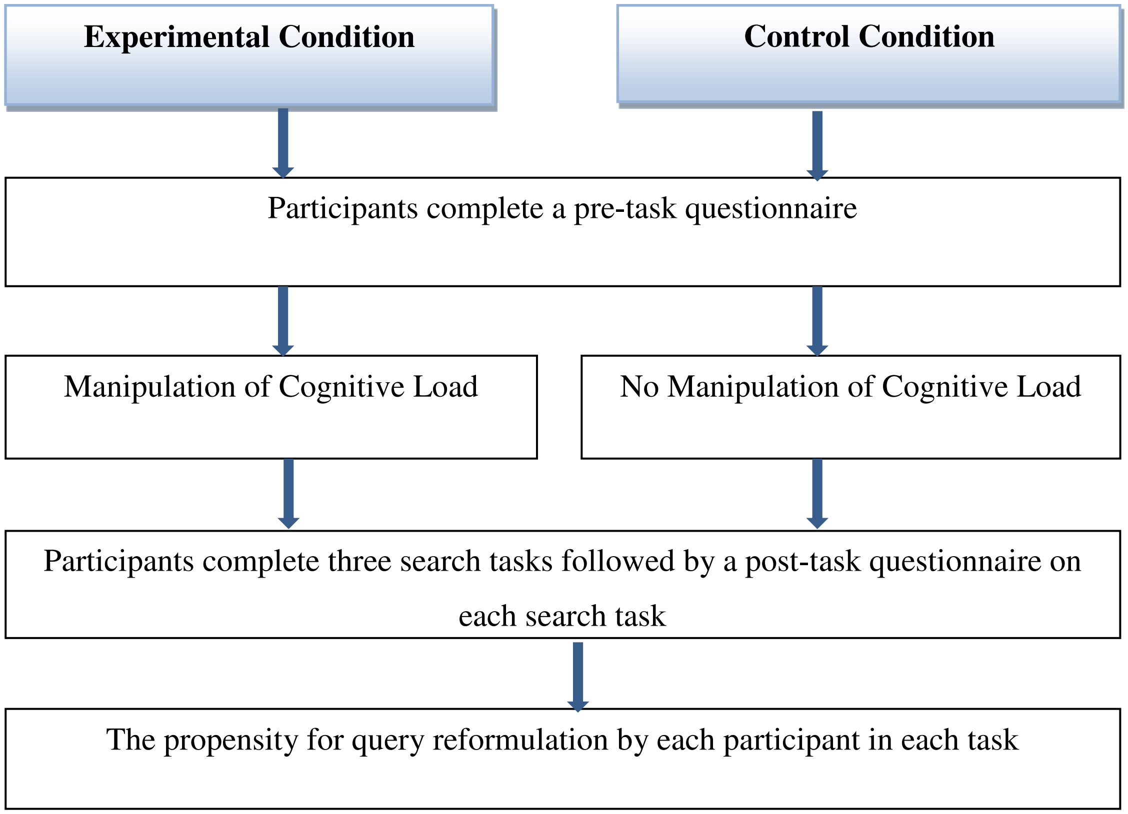

Figure 3.1. Conditions of Experiment

Subjects

The researcher recruited a total of 54 subjects from students enrolled at the School of Library and Information Studies at Florida State University (FSU). The target population was a group of university enrolled students. The gender profiles were well balanced consisting of 29 female and 25 male subjects. The subjects' birth year ranged from 1951 to 1992 (M = 1982, SD = 9.7). The degrees pursued consist of BA/BS (54%), MS/MA (13%), and Ph.D. (33%). The subjects included undergraduate and graduate students with varying levels of computer and Internet-searching skills. The researcher purposively recruited the subjects from the School of Library and Information Studies ranging from undergraduates to graduates because they are considered as experienced and trained searchers and may be familiar with terms used in this study such as information retrieval, precision and recall, or relevance.

Apparatus

A Dell OPTIPLEX 760 desktop computer (Microsoft Windows 7 Professional OS) was used to run the experiment in a laboratory. Camtasia 5.0 screen capture software and the Egg Timer Plus 3.12 application were installed to collect transaction data and lock the search task time interval.

Randomization of Subjects

The subjects were randomly assigned to one of the two groups through drawing of cards numbered from 1 – 54 upon arrival for the session. The subjects who drew 1-27 were assigned to the experimental group and the subjects who drew 28-54 were assigned to the control group. The card had a specified order of the experiments such as EG1A, EG2B, EG3C, CG1A, CG2B, etc. (See Table 3.2 & 3.3).

NASA-TLX Components

In this research, three of the six components of the National Aeronautical and Space Administration Task Load Index (NASA-TLX) – mental demand, temporal demand, and frustration—were manipulated in the experimental group. The full six components were then measured to obtain a subjective self-reported cognitive load score based on a weighted average of ratings as developed by Hart and Staveland (1988). NASA-TLX Index subjective rating scale has been used in previous research (Noyes & Bruneau, 2007) to assess subjective workload as an indication of cognitive load.

Manipulations: Mental Demand, Temporal Demand, and Frustration

The experiment consisted of three separate search tasks. Either mental demand, temporal demand, or frustration were manipulated during each search task for the experimental group. The order of the tasks was counterbalanced across the experimental group subjects. Block, Hancok, and Zakay (2010) state that different types of cognitive load manipulations provide a way to distinguish the independent variables used in experiments among various experimental conditions. This study classifies experimental manipulations of cognitive load according to six components of NASA-TLX workload assessment. The following components of NASA-TLX were adopted and modified to suit for this study.

1) Mental demand – A secondary task was added to the search task so that a dual-task demand was imposed on experimental group subjects to increase mental demand. The primary task wass a search task and the secondary task was to memorize an 8-digit number (i.e., all different 8 digits such as 324871659) and recall them following the completion of the search task. Pilot test subjects stated that they felt uncomfortable and had difficulty concentrating on the primary search task due to the burden of the secondary task, thus indicating increased mental demand.
2) Temporal demand – The experimental group was given a specific time duration in which to complete a search task in an attempt to increase temporal demand. Search tasks took 5-20 minutes during pilot tests. Kim and Rieh (2005) demonstrate that 15 minutes is adequate for most subjects to perform their usual search task behavior. The time for this task was therefore limited to 5 minutes to pressure subjects to complete the search task quickly. The Egg Timer Plus 3.12 application was be used to display a 5 minute countdown.
3) Frustration – The subjects were required to use an unfamiliar keyboard (the On-Screen Keyboard provided by Microsoft Windows 7 Professional) in an attempt to increase frustration during the task. Pilot test subjects expressed frustration and annoyance while using this keyboard to perform the task.

Search Tasks

Three ill-structured, problem-solving search tasks were developed to allow examination of cognitive activities while subjects performed each search task. The experiment consisted of three different search tasks that subjects performed in counterbalanced order (See Tables 3.2 and 3.3):

Virus Task

Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program, so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot log in to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Library Task

Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system such as library catalog system (not the library building). What are the best approaches to take and strategies to perform in this situation? Find the best strategies to rebuild the library computer system.

Web Security Task

Suppose you are a web developer who is in charge of preventing hacking and malicious attacks for a client's website. What kinds of prevention methods will you propose to your client?

Practice task

> Suppose you are a researcher exploring why people use social networking tools (i.e., twitter, facebook, youtube, blog, etc.)? What are the affective factors that cause people to join and use social networking tools listed above? What are advantages and disadvantages of using these social networking tools?

In this study, the search tasks are ill-structured, scenario-based, problem-solving search tasks requiring interaction between subjects and the system, i.e. user interaction scenarios. In each search task, subjects assumed that they have been hired in one of the three positions: information technology specialist, librarian, and web master. Rosoon and Carroll (2002) note that

> Representing the use of a system or application with a set of user interaction scenarios makes the system's use explicit, and in doing so orients design and analysis toward a broader view of computers. It can help designers and analysts to focus attention on assumptions about people and their tasks (p. 19).

Scenario-based design helps system designers to understanding and conceptualizing people's work and activities (Carroll, 1995) and she also notes that scenario-based design should be used to augment current system designs and enhance usability analysis in the system development lifecycle.

Counterbalancing the Sequence of the Search Tasks

The sequence of tasks has been shown to influence subjects' responses to search tasks (Fiske, Gilbert, & Lindzey, 2010; Goodwin, 2009). The sequence of tasks may have carryover effects from one task to the next. Therefore, in this study the tasks were counterbalanced across subjects using a Latin Square of order 3 * 3 arrays of ordered letters in which each letter appears once in each row and once in each column (see Table 3.2). The experiment involves three search tasks (Virus, Library, and Web Security). Each participant in each of the two groups performed all three tasks in counterbalanced order. Subjects in the experimental group also experienced three manipulations to cognitive load (i.e., mental demand, temporal demand, and frustration level). These manipulations appeared equally often in first, second, and third position of task order as shown in Table 3.2.

Table 3.2

Schematic representation of control of order effects by counterbalancing (Experimental Group)

Experimental Groups (n=27)	Task Orders	Task Sequence Available to Subjects with Manipulations			R
EG1 (n=9)	A (n=3)	(virus)(mental)	(library)(temporal)	(security)(frustration)	EG1A
	B (n=3)	(virus)(temporal)	(library)(frustration)	(security)(mental)	EG1B
	C (n=3)	(virus)(frustration)	(library)(mental)	(security)(temporal)	EG1C
EG2 (n=9)	A (n=3)	(library)(mental)	(security)(temporal)	(virus)(frustration)	EG2A
	B (n=3)	(library)(temporal)	(security)(frustration)	(virus)(mental)	EG2B
	C (n=3)	(library)(frustration)	(security)(mental)	(virus)(temporal)	EG2C
EG3(n=9)	A (n=3)	(security)(mental)	(virus)(temporal)	(library)(frustration)	EG3A
	B (n=3)	(security)(temporal)	(virus)(frustration)	(library)(mental)	EG3B
	C (n=3)	(security)(frustration)	(virus)(mental)	(library)(temporal)	EG3C

Note. * EG = Experimental Group, A, B, C = Task Order, R = Representation.

The purpose of counterbalancing the sequence of the search tasks was to have two factors (i.e., search tasks and manipulations) distributed equally and evenly to the subjects because there may be carryover effects. In the experimental condition, one-third of the experimental group (n=9) performed the virus search task with the mental demand, temporal demand, and frustration manipulation; one-third (n=9) performed the library search task with the mental demand, temporal demand, and frustration manipulation; and one-third (n=9) performed the security search task with the mental demand, temporal demand, and frustration manipulation.

More specifically, of 9 subjects from experimental group 1 (EG1), one-third (n=3) of subjects performed the virus search task with mental demand manipulation, the library search task with temporal demand manipulation, and the security search task with frustration manipulation (the order of EG1A); one-third (n=3) performed the virus search task with temporal demand manipulation, the library search task with frustration manipulation, and the security search task with mental demand manipulation (the order of EG1B); and one-third (n=3) performed the virus search task with frustration manipulation, the library search task with mental demand manipulation, and the security search task with temporal demand manipulation (the order of EG1C). Experimental group 2 (EG2) and group 3 (EG3) did a similar shuffling so that both search tasks and cognitive load manipulations are fully counterbalanced for order of appearance (see Table 3.2).

Thus, to counterbalance tasks this study required separate instructions for each of the sub-groups within the experimental group. Twenty-seven subjects were required for both the experimental and control group to ensure counterbalancing. The search tasks were balanced for the experimental group so that the order of cognitive load manipulation took place an equal number of times for each task (see Table 3.2).

In the control condition, 27 subjects were randomly assigned to order A, order B, and order C. One-third (n=9) completed Order A: virus task first, library task second, and security task third; one-third (n=9) completed Order B: library task first, security task second, and virus task third; and the remaining one-third (n=9) completed Order C: security task first, virus task second, and library task third (see Table 3.3).

Table 3.3

Schematic representation of control of order effects by counterbalancing (Control Group)

Control Groups (n=27)	Task Orders	Task Sequence Available to Subjects without Manipulations			R
CG1(n=9)	Order A	(virus)	(library)	(security)	CG1A
CG2(n=9)	Order B	(library)	(security)	(virus)	CG2B
CG3(n=9)	Order C	(security)	(virus)	(library)	CG3C

Note. * CG = Control group, A, B, C = Task Order, R = Representation.

Testbed

The Google search engine (http://www.google.com) was selected as the testbed for this research because it is well-known, widely used, and indexes billions of web pages (Wikipedia, 2010). Searchers not only perceive it to be simple and easy to use, but also expect a high possibility of relevant search results. It has a simple search interface that provides both basic and advanced search options. Figure 3.2 below shows the current user interface for Google search engine.

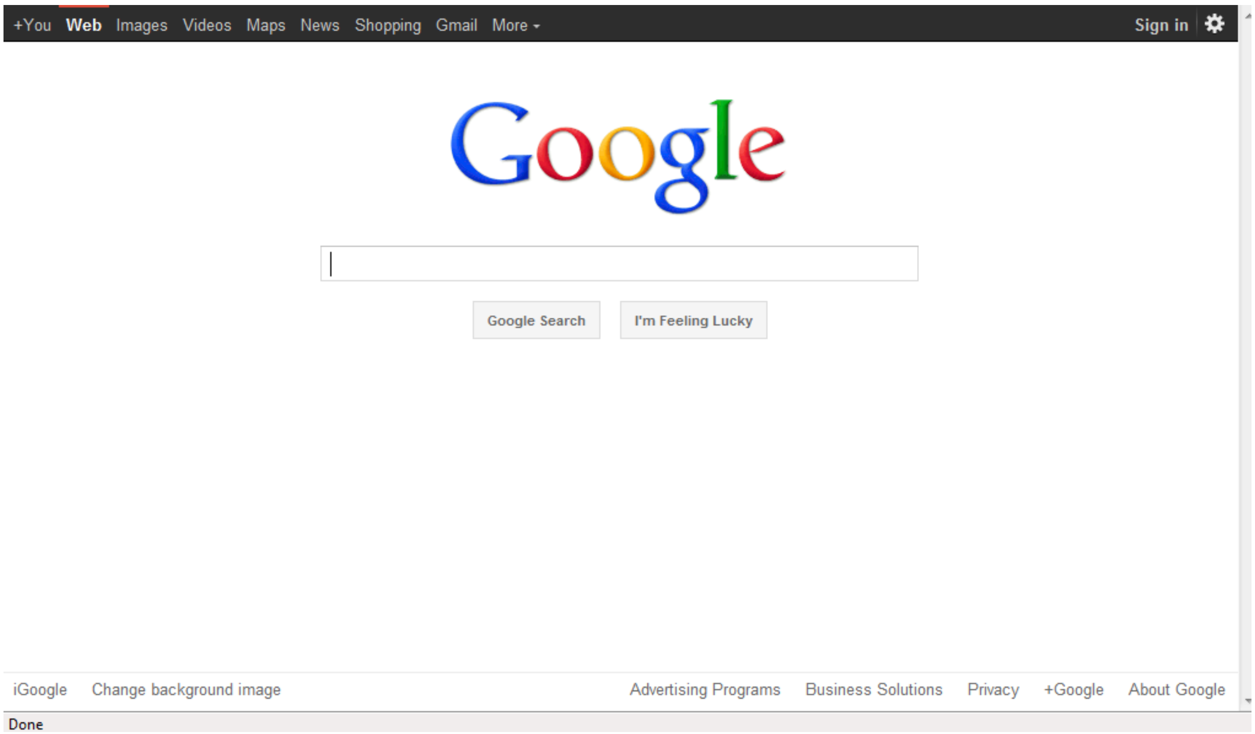

Figure 3.2. Current user interface of Google search engine (Captured 1/05/2012)

Experiment Instruments

Surveys

Two survey questionnaires were administered: a pre-task survey and a post-task survey. The pre-task survey collected demographic data and information about previous experience including computer, web-based searching, and Google search engine experience. It was administered to each participant at the beginning of the experiment session. The pre-survey includes two sections: Section 1--Demographics and Section 2--Previous Experience (See Appendix A)

At the end of each search-task a post-task survey questionnaire collected data about subjects' experiences with cognitive load during the search task and query reformulation. Since the experiment included three search tasks, each participant completed three post-task questionnaires. Questions about cognitive load were developed from the NASA-TLX workload assessment and elicited responses about mental demand, physical demand, temporal demand, performance, effort, and frustration. Questions about query reformulation solicited responses about when, why, and in what ways query reformulation was needed. The post-task survey questionnaire includes two sections: Section 1--Mental demand, physical demand, temporal demand, performance, effort, and frustration across the search process (developed and modified from the NASA-TLX workload assessment); Section 2--Query reformulation experience (See Appendix B)

Search Log Transaction

Transaction logs of each search were recorded using Camtasia software on the researcher's computer. These search logs were examined to see if there are any differences in the propensity for query reformulation behavior between the experimental and control groups.

Data Preparation for Search Logs

Recorded search log data were imported into SPSS software. Search log data could then be parsed into entities, focusing user IDs, terms, phrases, queries, sessions, and more as shown in Table 3.4.

Table 3.4

Entity Schema (Unit of Analysis)

Entity name	Description
userid	User identification
sdate	Session date occurred
stime	Session time used
search_urls	Search URLs visited

Table 3.4 - continued

Entity name	Description
total_num_doc	Total number of documents viewed
total_num_query	Total number of queries used
total_num_qeury reformuation	Total number of query reformulation used
num_clicks	Total number of clicks
length_term_query1	Total number of query terms used in the initial query
length_term_query2	Total number of query terms used in the second query
length_term_query3	Total number of query terms used in the third query
length_term_query4	Total number of query terms used in the fourth query
…	…
length_term_queryN	Total number of query terms used in the Nth query

Data Coordination and Interpretation

Data were collected, coordinated, and interpreted through several methods (see Table 3.5):

- Pre-survey questionnaire--subjects were instructed to fill out a pre-survey.
- Search transaction logs in experiments--subjects were instructed to perform three search tasks. Screen-captures of the information search processes were made using Camtasia software.
- Post-survey questionnaire--subjects were asked to fill out a post-survey at the end of each search task rating their cognitive load.

Table 3.5

Data Coordination and Interpretation

Phases	Data	Interpretation
Pre-survey	Demographic, Computer & Web search experience, Google search experience, and Google search query refinement experience	Data interpreted and comparisons made between the experimental group and the control group using statistical software (SPSS) to derive frequencies, means, and analyses of variance (ANOVA).
Post-survey	Self-reported subjective Cognitive Load rating from NASA-TLX workload assessment	Data interpreted and the mean Likert scale rating for each component of cognitive load compared between the two groups.
Search transaction Logs	Recorded Search Log on propensity for query reformulation	Data interpreted and comparisons made between the two groups using statistical software (SPSS) to derive frequencies, means, t-test, and analyses of variance (ANOVA).

Reliability and Validity

Reliability is the degree to which questions within a survey or trials within an experiment measure the same construct over time or the degree to which questions deliver similar responses over time (Creswell, 2003; Creswell & Plano-Clark, 2007). The researcher believes that if another researcher performs this experiment exactly as described with similar questions and under similar conditions, similar results will be achieved; under those circumstances, the results would have high reliability.

Construct validity, internal validity, and external validity are also measures of the quality of experimental results. Construct validity is the extent to which a variable reflects the theoretical construct that the study intends to measure (Creswell, 2003; Creswell & Plano-Clark, 2007). Since this experiment is designed to inform about the effects of cognitive load on propensity for query reformulation, the cognitive load of the experimental group must be successfully manipulated for there to be high construct validity. The technique this study uses to measure the propensity for query reformulation must also produce a true measure of the propensity for query reformulation for there to be high construct validity.

To satisfy the first requirement, the variables (manipulations) must be an indication of theoretical constructs whose cause and effect relationship this study is trying to investigate. The manipulations were checked to see that they performed as expected. That is, effect of the manipulations on the mental demand, temporal demand, and frustration level by measuring the NASA-TLX mental workload rating of to see if the experimental group had a level higher than that found in the control group. The researcher also checked whether or not these manipulations affected the propensity for query reformulation (second requirement) by comparing the number of query reformulation attempts performed by the experimental group to the number of attempts performed by the control group.

Although the experiment aimed to explore the effects of cognitive load on the propensity for query reformulation, it was also necessary to consider the possibility that other variables might influence the propensity for query reformulation. The way subjects perform in the search task can be influenced not only by the level of cognitive load, but also by other variables such as time of day, anxiety, tiredness, the amount of sleep the subjects got the previous night, etc. The researcher endeavored to hold these nuisance or extraneous variables constant or make them random. Control techniques such as random selection and allocation were incorporated into the experimental design to handle these variables. Counterbalancing techniques were also employed since order effects may occur. The researcher instructed the subjects to select a time of day that ensured that they would have had sufficient sleep and would not be tired. The subjects were told that the experiment would not try to measure how quickly they completed each search, but rather that they would be advised to feel comfortable during the experiment. A statistical validity check was employed to see if manipulations would work as they are supposed to by increasing with each manipulation.

External validity is the extent to which the relationship between the variables observed by the researcher in the context of the experiment can be generalized to different contexts and individuals (Creswell, 2003; Creswell & Plano-Clark, 2007). External validity may be subdivided into three specific types: ecological, population, and temporal validity (Brewer &

Hunter, 2005). The researcher should consider whether the findings can be generalized to settings other than that of this study, to people who differ in some important respects from those who participate in this experiment, and to other time periods.

A control group and randomization of subjects between the groups were used in this experiment to reduce internal validity issues. This study employed a laboratory experiment design allowing the researcher to randomly assign subjects to the two groups so that threats to internal validity are minimized. Furthermore, the laboratory was a setting providing the experimenter greater control over potential causal variables.

Summary

This chapter described the exploratory experimental design that was employed in this research. It also discussed data collection and analysis techniques. A full description of the data analysis will be provided in chapter four.

CHAPTER FOUR

RESULTS AND DISCUSSION

The purpose of this exploratory study is to determine if there are any differences in the propensity for query reformulation behavior between searchers who experience cognitive load manipulation and searchers who do not experience cognitive load manipulation. This chapter presents statistical data analysis and results. The methods employed consist of a pre-survey, an experiment (three search tasks), and a post-survey. In the pre-survey questionnaire, demographics and experience using computers, searching, and specifically Google search experience were analyzed. The search task experiment was conducted to determine whether or not there were differences in the propensity for query reformulation behavior between the experimental group and control group. In the post-survey questionnaire, NASA-TLX workload assessment was analyzed to see if there were any differences in cognitive load between the two groups corresponding to three manipulations: mental demand, temporal demand, and frustration.

Assumptions

This study assumes that there are differences in the cognitive load for searchers who are exposed to manipulations and searchers who are not exposed to manipulations. More importantly, this study assumes that there are differences in the propensity of query reformulation behavior between searchers who are exposed to cognitive load manipulation and searchers who are not exposed to cognitive load manipulation when searching with Google .

Results

Characteristics of Subjects

A total of 54 subjects participated in this study; all of whom were either undergraduate or graduate students enrolled in the College of Communication and Information at Florida State University. The gender profiles were well balanced consisting of 29 female and 25 male subjects. Subjects were randomly assigned to either an experimental or control group. Subjects' computer knowledge and web search skills were above average as was to be expected since they were enrolled in educational programs that require these skills.

The pre-survey collected information on gender, birth year, and degree sought. As shown in table 4.1, gender is adequately distributed with 54% female and 46% male. In the control group, gender is also adequately distributed with 44% in female and 56% male. The female portion sampled in the experimental group is slightly larger than the male with 63% female and 37% male. The subjects' birth year ranged from 1951 to 1992 (M = 1982, SD = 9.7). The degrees pursued consist of BA/BS (54%), MS/MA (13%), and Ph.D. (33%) as shown in table 4.1. None of the demographic variables yielded significant differences between the experimental group and the control group regarding the research questions in this study.

Table 4.1

Summary for Gender, Age, and Degree Sought

Variables	Range	Group		All
		CG (n=27)	EG (n=27)	Total (N=54)
Gender	Male	15 (55.6%)	10 (37.0%)	25 (46.3%)
	Female	12 (44.4%)	17 (63.0%)	29 (53.7%)
Age	Less than 20	1 (3.7%)	0 (0%)	1 (1.8%)
	21-30	16 (59.3%)	16 (59.3%)	32 (59.3%)
	31-40	6 (22.2%)	7 (25.9%)	13 (24.1%)
	Over 41	4 (14.8%)	4 (14.8%)	8 (14.8%)
Degree sought	BA/BS	14 (51.9%)	15 (55.6%)	29 (53.7%)
	MS/MA	5 (18.5%)	2 (7.4%)	7 (13.0%)
	Ph.D.	8 (29.6%)	10 (37.0%)	18 (33.3%)

Note. CG = Control Group; EG = Experimental Group.

Computer Knowledge

When subjects were asked about computer knowledge, 17 out of 54 subjects rated their computer knowledge as average (31.5%), 29 out of 54 subjects rated it as high (53.7%), and 8 out of 54 subjects rated it as very high (14.8%). 68.5% of the subjects answered their computer knowledge as high or very high.

Web Search Knowledge

Subjects rated their web search knowledge as at or above average. Twelve out of 54 subjects rated it as average (22.2%), 28 subjects rated it as high (51.9%), and 14 subjects rated it as very high (25.9%).

Web Search Skills

Subjects rated their web search skills (e.g., Boolean language, advanced search, etc.) as at or above average. Fourteen out of 54 subjects rated it as average (25.9%), 30 rated it high (55.6%), and 9 rated it very high (16.7%).

Google Search Engine Use Frequency

It is interesting to note that no subjects selected "more than once a day" concerning their Google use frequency. However, 55.6% (30/54) reported that they use Google for their daily information seeking on the web. The Google use frequency dataset divided the subjects' responses into "seldom," "weekly," "at least once a week," "daily," and "more than once a day." It is interesting to see more than half of the subjects use the Google search engine. It might have been useful to also ask what other search engines they use for their daily information seeking. Likewise, it might have been helpful to ask the exact number of times per day they use Google for information seeking.

Google Search Query Refinement

Twelve out of 54 subjects rated their Google search query refinement at 1 time (22%), 19 subjects rated it about 2 times (35.2%), and 19 subjects rated it more than 3 times (35.2%). It is interesting to see that 50 out of 54 subjects (92.6%) involved in making search query refinement more than one time (>=1 query) when they used the Google search engine. The Google search query refinement dataset divided the subjects' responses into "never," "rarely," "about 1 time," "about 2 times," and "more than 3 times."

Google Search Experience

When the subjects were asked about Google search experience, it is interesting to see that there were no subjects who selected "more than once a day" as their rating; 55.6% (30/54) of the subjects reported that they seldom use Google search engine for their daily life information seeking on the web. The Google search experience data set divided the subjects' responses into "never," "rarely," "sometimes," "usually," and "always." It is interesting to see that more than half of the subjects seldom use Google search engine. It might be useful in future research to ask what other search engines are used for their daily life information seeking on the web.

As shown in Figure 4.1 below, it indicates that there are no significant differences between the control group and the experimental group in terms of computer knowledge, Web search knowledge, Web search skills, Google search experience, Google search query refinement, and Google search success. The subjects' self-reported experience on a Likert scale from 1 "very low" to 5 "very high" concerning their computer knowledge, web search knowledge, and web search skills; from "seldom" to "more than once a day" for Google use frequency; from "never" to "more than 3 times per day" for Google query refinement frequency; and from "never" to "always" for Google search success.

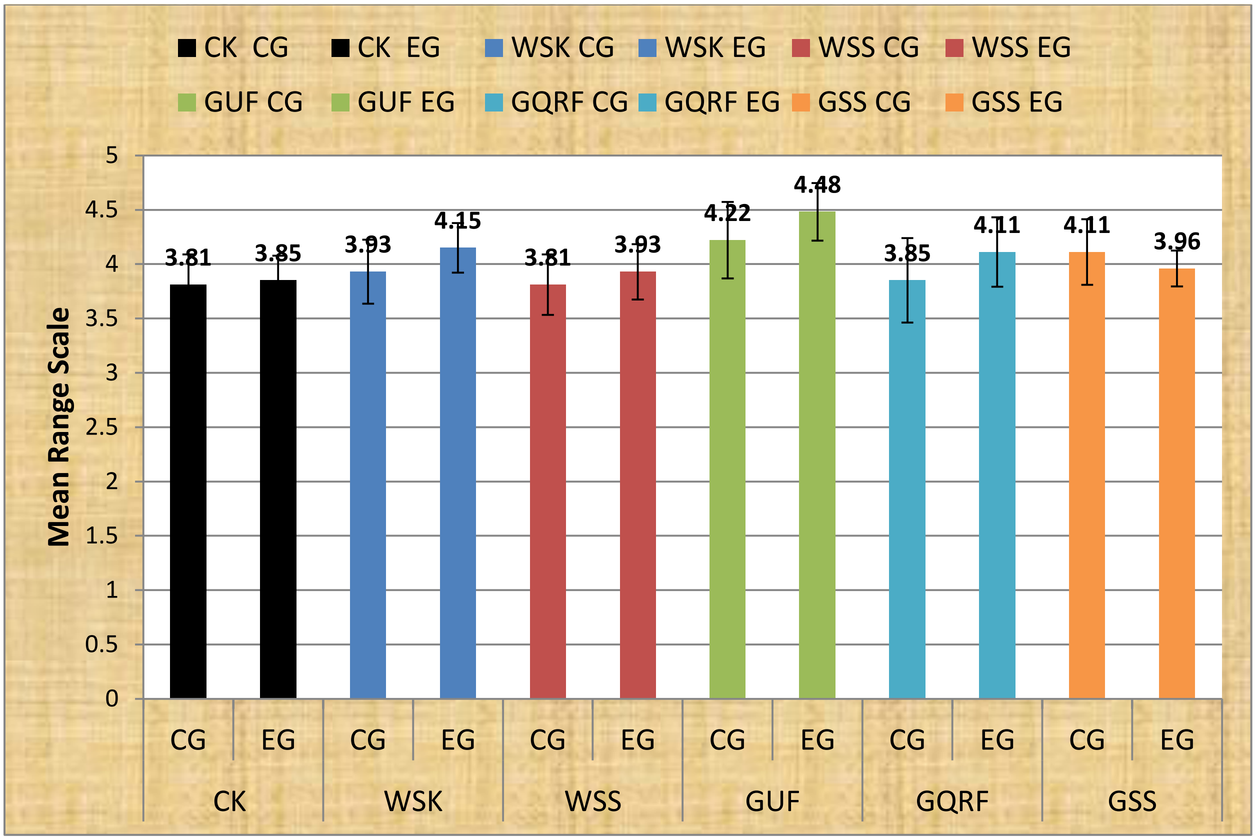

Figure 4.1. Comparisons between the control and the experimental group in computer knowledge, web search knowledge, web search skills, Google use frequency, Google query refinement frequency, and Google search success. CG = Control Group; EG = Experimental Group, X-axis of CK = Computer knowledge; WSK = Web Search Knowledge; WSS = Web Search Skills; GUF = Google Use Frequency; GQRF = Google Query Refinement Frequency; GSS = Google Search Success. Y-axis of CK, WSK, WSS: ranging from 1 (very low), 2 (low), 3 (average), 4 (high), to 5 (very high). Y-axis of GUF: ranging from 1 (seldom), 2 (weekly), 3 (at least once a week), 4 (daily), to 5 (more than once a day). Y-axis of GQRF: ranging from 1 (never), 2 (rarely), 3 (about 1 time), 4 (about 2 times), to 5 (more than 3 times). Y-axis of GSS: ranging from 1 (never), 2 (rarely), 3 (sometimes), 4 (usually), to 5 (always).
Error bars show the 95% confidence intervals for the mean.

Search Tasks

This study used a within-subjects design with repeated measures ANOVA and t-test analysis. The experiment was set up with three manipulations within the experimental group (i.e., mental demand, temporal demand, and frustration) using a dual-task involving a memory load, time constraint, an on-screen keyboard time lock, and a frustration to increase data entry difficulty. The subjects' were asked to find relevant information for three search tasks (i.e., a virus search task, a library search task, and a security search task). Therefore, manipulation served as an independent variable for the three different manipulations of mental demand, temporal demand, and frustration. The dependent variables for this experiment were the propensity of query reformulation and three manipulations of NASA-TLX that contributed to cognitive load. Furthermore, six components of subjective workload from the NASA-TLX score were assessed to examine if there were differences in the score of NASA-TLX between the experimental and control group.

Table 4.2 below summarizes the query reformulations in the control and experimental groups including mean, median, standard deviation, variance, range, minimum to maximum, and percentiles. The mean values of query reformulations for the experimental group are significantly smaller than those of the control group (i.e., 1.89 < 4.37, 2.04 < 4.48, and 2.41 < 4.52). On average, query reformulation was reduced to half the normal level when the cognitive load was increased. The value corresponding to the 75th percentile was 6 query reformulations in task1, 6 in task2, and 9 in task3 for the control group, while the value corresponding to the 75th percentile for the experimental group is 2 query reformulations in task1, 3 in task2, and 4 in task3.

Table 4.2

Summary for the propensity of query reformulation in control group and Experimental group

	Control Group			Experimental Group		
	QR in Task1	QR in Task2	QR in Task3	QR in Task1	QR in Task2	QR in Task3
N (valid)	27	27	27	27	27	27
Mean	**4.37**	**4.48**	**4.52**	**1.89**	**2.04**	**2.41**
Median	4.00	3.00	3.00	1.00	1.00	2.00
Std. Deviation	2.748	2.992	3.735	1.717	2.295	2.188
Variance	7.550	8.952	13.952	2.949	5.268	4.789
Range	10	11	11	7	11	8
Minimum	1	1	0	0	0	0

Table 4.2 - continued

	Control Group			Experimental Group		
	QR in Task1	QR in Task2		QR in Task1	QR in Task2	
Maximum	11	12	11	7	11	8
Percentiles 25	3.00	2.00	2.00	1.00	1.00	1.00
Percentiles 50	4.00	3.00	3.00	1.00	1.00	2.00
Percentiles 75	6.00	6.00	9.00	2.00	3.00	4.00

Note. Task1 = virus search task; Task2 = library search task; Task3 = security search task.

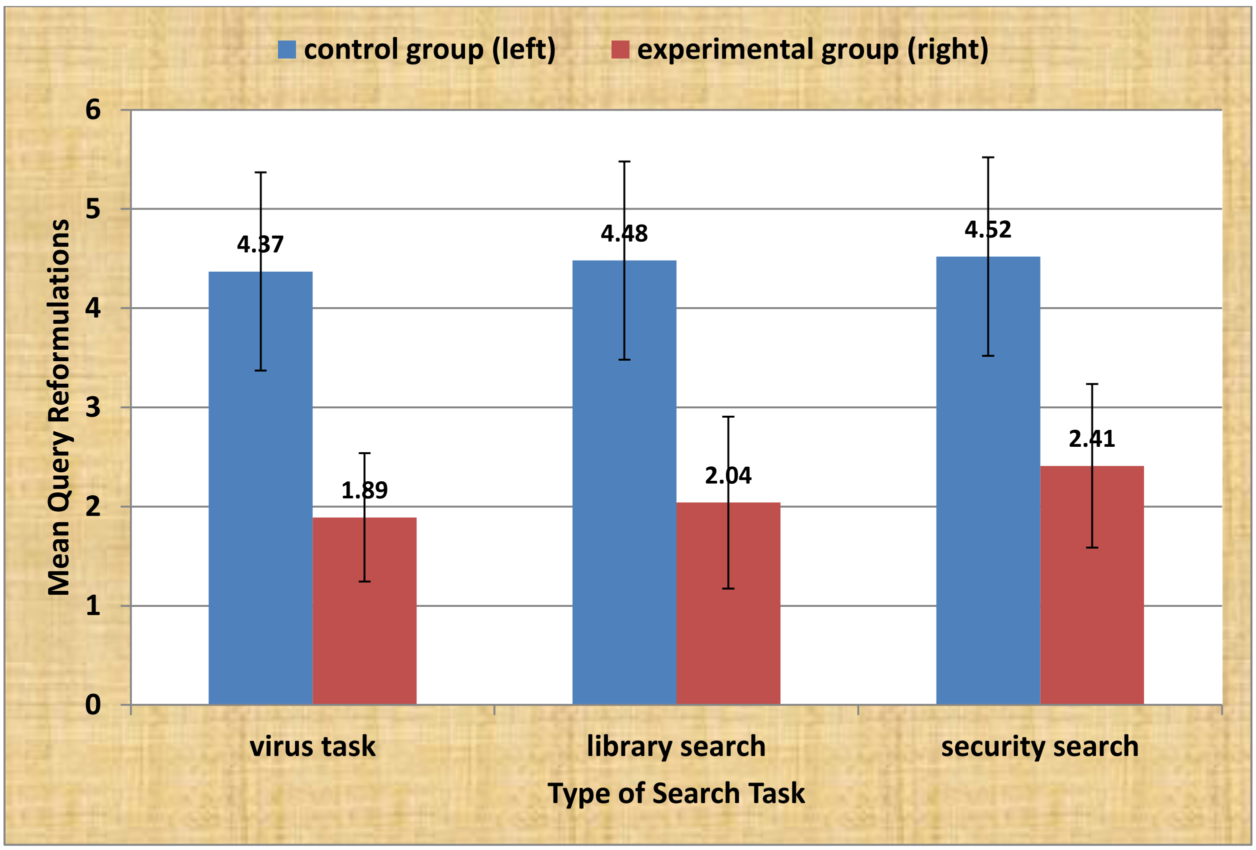

Figure 4.2. Mean query reformulation propensity with the type of search tasks. Error bars show the 95% confidence intervals for the mean.

As displayed in Figure 4.2, it is clear that estimated marginal means of query reformulations in three search tasks significantly differ between the control group and the experimental group, and there is no significant difference in the propensity for query reformulation in terms of the type of search tasks.

There were three search tasks given to subjects in this study: a virus search task, library search task, and security search task. These tasks were given to both the experimental and control group. However, cognitive load manipulations were only added to these search tasks in the experimental group.

As shown in pairwise comparison on the type of search tasks (Table 4.3), there are no statistically significant interactions between the types of search tasks on the propensity for query reformulation. By looking at the significance values and the means, the results show that the propensity to reformulate queries was not significantly affected by the types of search tasks given ($p = 1.000$) at $\alpha = 0.05$ level. The tasks do not differ at $\alpha = 0.05$ level (95% confidence level) with $p = 1.000$ to the propensity for query reformulation in this study. Therefore, the propensity for query reformulation in each of three search tasks was focused and analyzed in this study for three-sub research questions.

Table 4.3

Pairwise Comparison on the Type of Search Tasks in the Propensity of Query Reformulation

(I) tasks	(J) tasks	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-.130	.476	**1.000**	-1.307	1.048
	3	-.296	.476	**1.000**	-1.474	.882
2	1	.130	.476	**1.000**	-1.048	1.307
	3	-.167	.467	**1.000**	-1.323	.989
3	1	.296	.476	**1.000**	-.882	1.474
	2	.167	.467	**1.000**	-.989	1.323

Note. 1=virus search task; 2=library search task; 3=security search task.

Discussion

The cognitive load was measured with a subjective NASA-TLX questionnaire and the propensity of query reformulations was examined in the searchers' transaction logs that were screen-captured. In NASA-TLX, three components of cognitive load (mental demand, temporal demand, and frustration) were also examined between experimental and control group for manipulations.

Key Research Question

Is there any difference in the propensity for query reformulation behavior between searchers who experience cognitive load manipulation and searchers who do not experience cognitive load manipulation?

An ANOVA was employed for the key research question. There was a statistically significant difference between control and experimental group on the propensity of query reformulations with (sig. < .05, 95% confidence) p= .000 in virus search task; p= .001 in library search task; and p=.014 in security search task (Table 4.4). The ANOVA table below tests whether the difference between groups is significantly higher than the deviations within each group. The Sig. value indicates that the between group's variation can explain a relatively small portion of the variation in the propensity of query reformulation. As such, it makes sense to go further and compare the difference in the mean of the propensity of query reformulations across the three manipulations.

Table 4.4

Summary for Descriptive Statistics for Query Reformulation in Each Task with ANOVA

		N	Mean	SD	SE	95% CI for Mean		Min	Max	F	Sig.
						LB	UB				
QR t1	CG	27	**4.37**	2.748	.529	3.28	5.46	1	11	15.836	**.000**
	EG	27	**1.89**	1.717	.330	1.21	2.57	0	7		
	Total	54	**3.13**	2.592	.353	2.42	3.84	0	11		
QR t2	CG	27	**4.48**	2.992	.576	3.30	5.67	1	12	11.346	**.001**
	EG	27	**2.04**	2.295	.442	1.13	2.94	0	11		
	Total	54	**3.26**	2.915	.397	2.46	4.05	0	12		
QR t3	CG	27	**4.52**	3.735	.719	3.04	6.00	0	11	6.421	**.014**
	EG	27	**2.41**	2.188	.421	1.54	3.27	0	8		

Table 4.4 - continued

		N	Mean	SD	SE	95% CI for Mean		Min	Max	F	Sig.
						LB	UB				
	Total	54	**3.46**	3.214	.437	2.59	4.34	0	11		

Note. CG = Control Group; EG = Experimental Group, QR t1 = total Query Reformulation in task1; QR t2 = total Query Reformulation in task2; QR t3 = total Query Reformulation in task3, SD = Std. Deviation; SE = Std. Error; CI = Confidence Interval; LB = Lower Bound; UB = Upper Bound; Min = Minimum; Max = Maximum; ANOVA in 95% confidence interval.

The descriptive summary of query reformulation in Table 4.4 above indicates that the average number of query reformulations for the control group during the virus search task was $M = 4.37$, $SD = 2.748$ with a 95% confidence that the real average would fall between 3.28 and 5.46, while for experimental group the virus search task was 1.89 with a 95% confidence that the real average would fall between 1.21 and 2.57. Similarly, in the library search task the average number of query reformulations for control group was $M = 4.48$, $SD = 2.992$ with a 95% confidence that the real average would fall between 3.30 and 5.67, while the experimental group was 2.04 with a real average that falls between 1.13 and 2.94. Lastly, in security search task the average number of query reformulations for control group was $M = 4.52$, $SD = 3,735$ and the real average would fall between 3.04 and 6.00, while the experimental group was 2.41 and the real average would fall between 1.54 and 3.27.

On average, regardless of manipulations, the total number of query reformulations was 3.13 in the virus task, 3.26 in the library task, and 3.46 in the security task (Table 4.4). Based on these data, one can see that the average number of query reformulations is around 3 or 4 queries. The results of this study support findings by Jansen et al. (2005) showing that the proportion of users who modified queries was 52%, with 32% making 3 or more queries within the session.

Those exposed to cognitive load manipulations (i.e, experimental group), namely, *mental demand*, *temporal demand*, and *frustration*, made 2.18 times fewer search queries on average than searchers not exposed (i.e., control group) (See Figure 4.3 & Table 4.5).

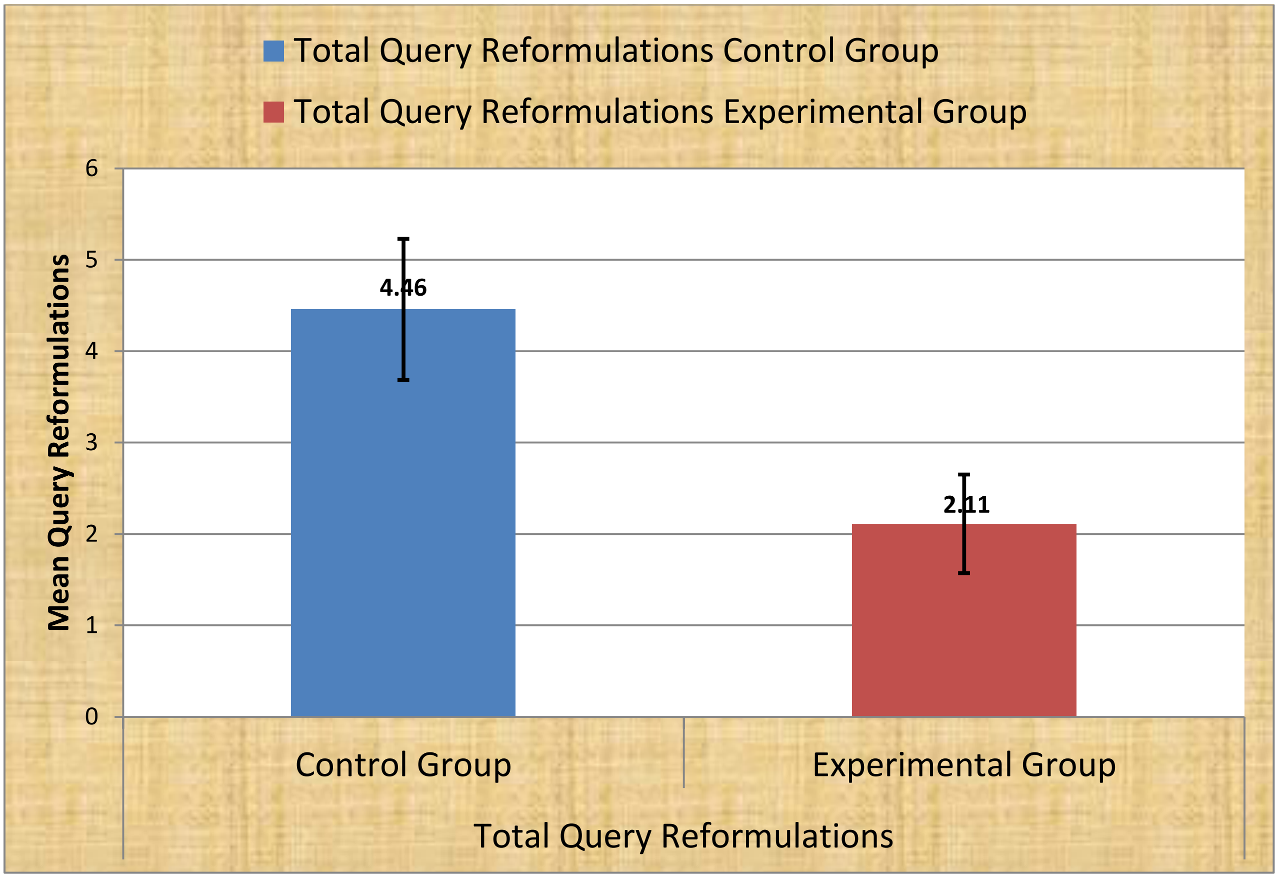

Figure 4.3. Total query reformulations by groups. Error bars show the 95% confidence intervals for the mean.

Table 4.5

Total Query Reformulations Comparison between Experimental Group and Control Group for All Three Search Tasks in Total with t-test

	Group	N	M	MD	SD	SEM	SED	t	df	Sig. (2-tailed)	95% CI	
											LB	UB
TQR	CG	27	4.46	2.35	2.04469	.39350	.48010	4.886	52	**.000**	1.38228	3.30908
	EG	27	2.11		1.42924	.27506						

Note. TQR = Total number of Query Reformulations, CG = Control Group; EG = Experimental Group, M = Mean; MD = Mean Difference; SD = Std. Deviation; SEM = Std. Error Mean; SED = Std. Error Difference; LB = Lower Bound; UB = Upper Bound.

As shown in Table 4.6 below , pairwise comparisons also provide evidence that the propensity for query reformulation between two the groups is statistically significant at the $p<.05$ level with a 2.321 mean difference and a .479 standard error.

Table 4.6

Summary for Pairwise Comparisons, Mean Difference, Standard Deviation, and Significance of Query Reformulation between Groups

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
CG	EG	2.321*	.479	**.000**	1.360	3.282
EG	CG	-2.321*	.479	**.000**	-3.282	-1.360

Note. CG = Control Group; EG = Experimental Group, i.e., manipulations added (experimental group), no manipulations added (control group), Based on estimated marginal means
*. The mean difference is significant at the .05 level.
a. Adjustment for multiple comparisons: Bonferroni.

Table 4.7

Query Reformulation in Virus Search Task with or without Manipulations

Manipulations	Mean	Std. Deviation	Std. Error of Mean	Minimum	Maximum	Range
No manipulation	4.37	2.748	.529	1	11	10
Mental demand	2.89	2.369	.790	0	7	7
Temporal demand	1.22	.972	.324	0	3	3
Frustration	1.56	1.130	.377	0	4	4
Total	3.13	2.592	.353	0	11	11

Table 4.7above shows that a significant difference in the propensity for query reformulation behavior in the virus search task was found between searchers who were exposed to cognitive load manipulations and searchers who were not exposed. Those exposed to the manipulations of mental demand (2.89), temporal demand (1.22), and frustration (1.56) made fewer search queries than searchers who were not exposed (4.37). It is interesting to see that the maximum number of query reformulations for searchers who were not exposed to cognitive load manipulations is 11 ranging 1 to 11. On the other hand, those exposed to the manipulations of mental demand made 7 maximum number of query reformulations, those exposed to the

manipulations of temporal demand made 3 maximum number of query reformulations, and those exposed to the manipulations of frustration made 4 maximum number of query reformulations.

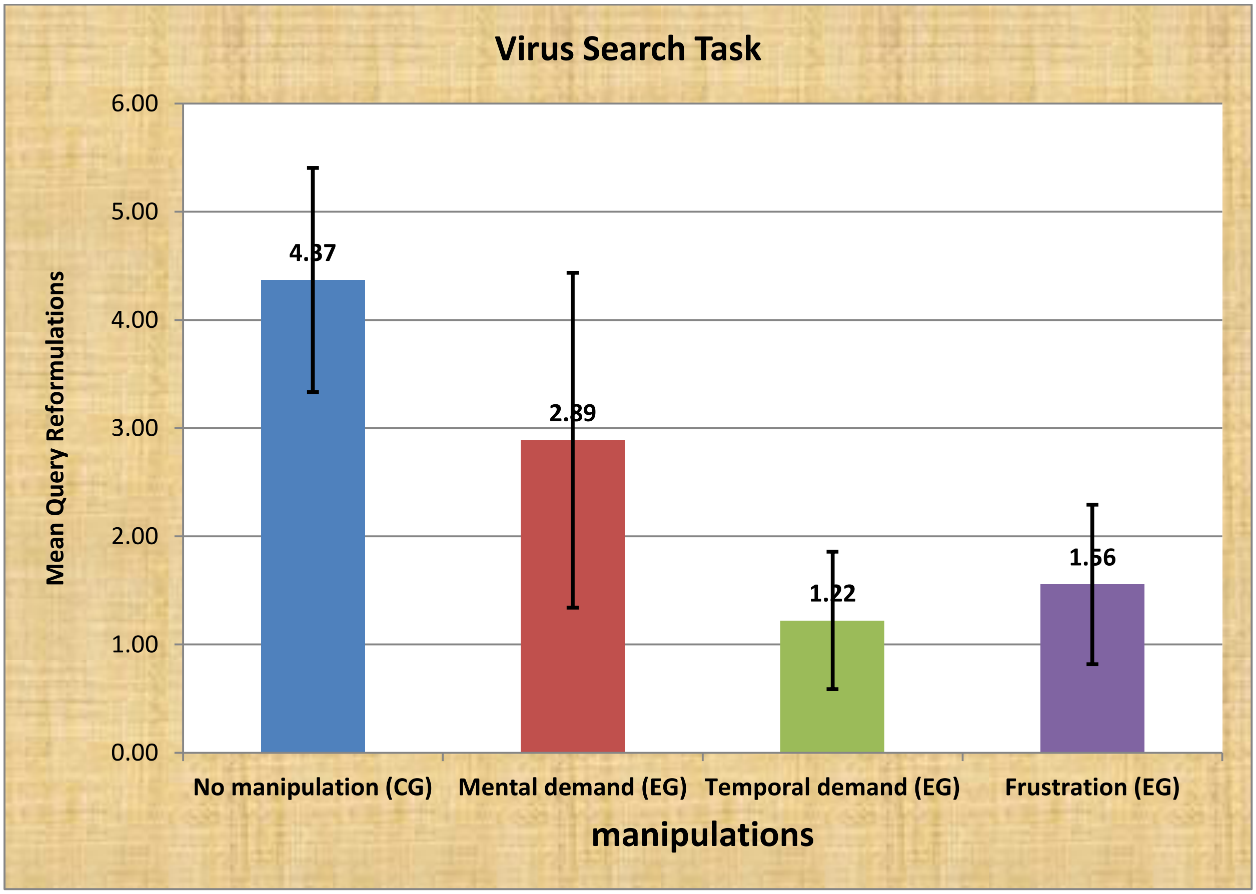

Figure 4.4. Virus search task. Error bars show the 95% confidence intervals for the mean.

Table 4.8

Query Reformulation in Library Search Task with or without Manipulations

Manipulations	Mean	Std. Deviation	Std. Error of Mean	Minimum	Maximum	Range
No manipulation	4.48	2.992	.576	1	12	11
Mental demand	2.33	3.464	1.155	0	11	11
Temporal demand	2.11	1.453	.484	0	5	5
Frustration	1.67	1.658	.553	1	6	5
Total	3.26	2.915	.397	0	12	12

A significant difference in the propensity for query reformulation behavior in the library search task was found between searchers who were exposed to cognitive load manipulations and searchers who were not exposed (Table 4.8). Those exposed to mental demand (2.33), temporal demand (2.11), and frustration (1.67) made fewer search queries than searchers who were not exposed (4.48).

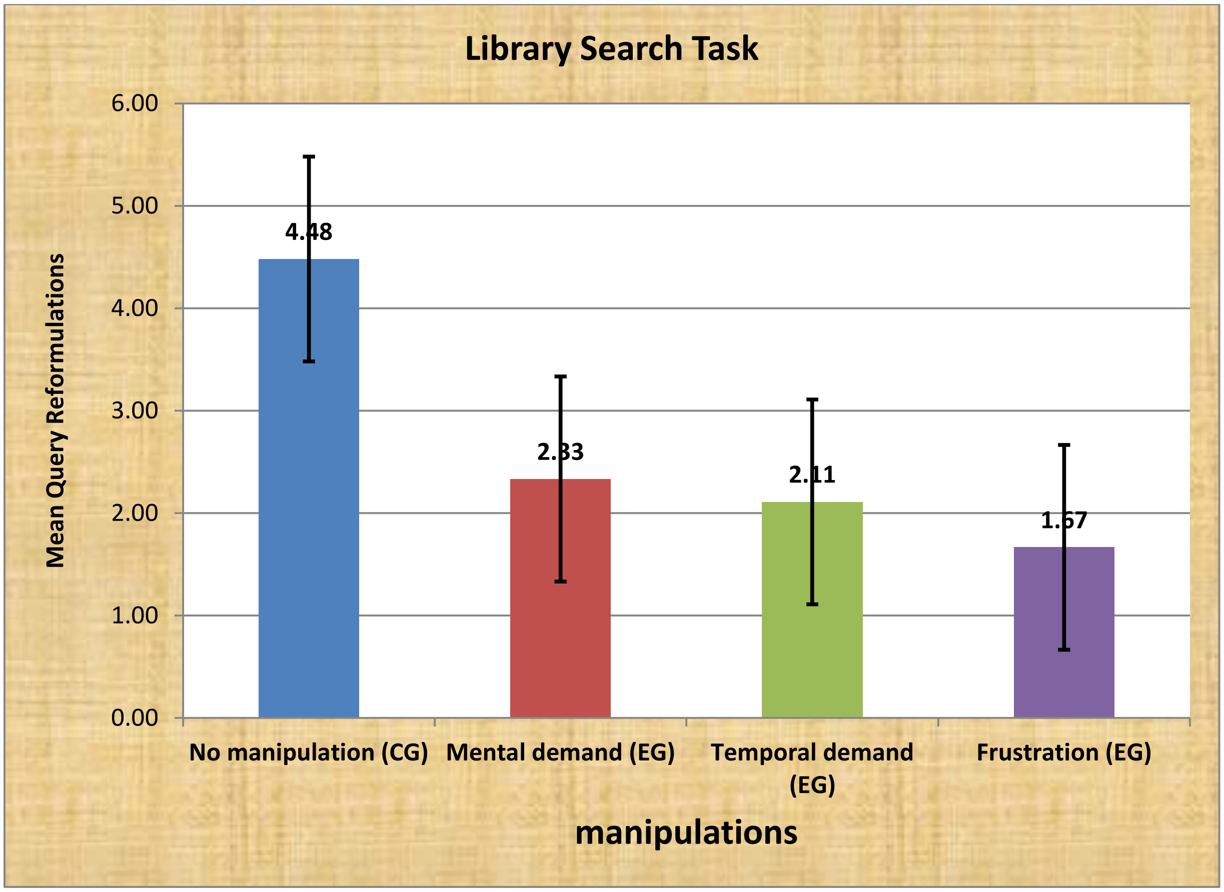

Figure 4.5. Library search task. Error bars show the 95% confidence intervals for the mean.

Table 4.9

Query Reformulation in Security Search Task with or without Manipulations

Manipulations	Mean	Std. Deviation	Std. Error of Mean	Minimum	Maximum	Range
No manipulation	4.52	3.735	.719	0	11	11
Mental demand	3.11	3.180	1.060	0	8	8
Temporal demand	2.00	1.225	.408	0	4	4
Frustration	2.11	1.764	.588	0	5	5
Total	3.46	3.214	.437	0	11	11

As shown in Table 4.9, in the security search task a significant difference in the propensity for query reformulation behavior was found between searchers who were exposed to cognitive load manipulations and searchers who were not exposed. Those exposed to mental demand (3.11), temporal demand (2.00), or frustration (2.11), made fewer search queries than searchers who were not exposed (4.52).

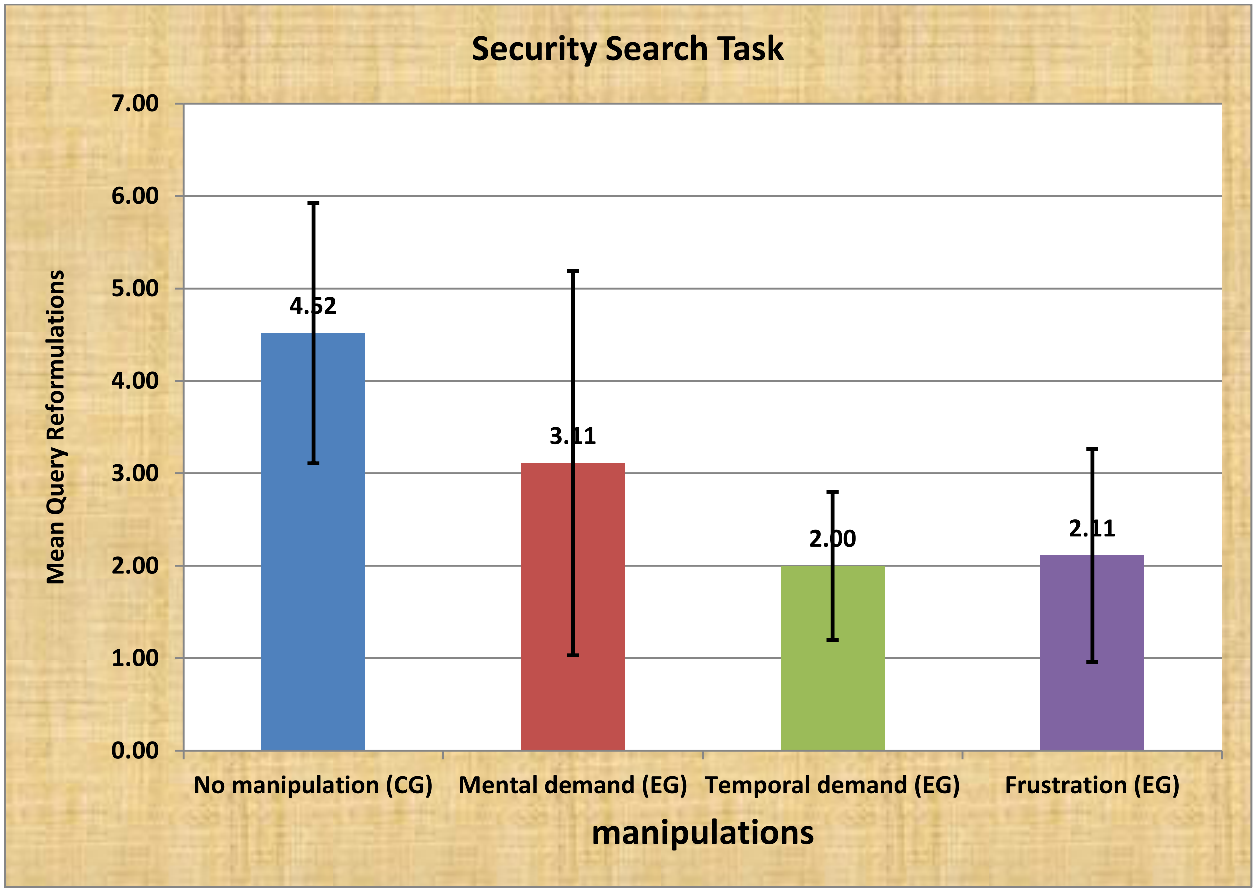

Figure 4.6. Security search task. Error bars show the 95% confidence intervals for the mean.

Sub-research question 1(Mental Demand Manipulation)

Is there any difference in the propensity for query reformulation behavior when mental demand increases?

For this question, a t-test was employed (Table 4.10 below). As discussed earlier, the propensity for query reformulation between groups did not differ with respect to the types of tasks. Thus, in this analysis, the three search tasks were combined for each of the three sub-research questions. In the case of the mental demand manipulation, the propensity for query reformulation with respect to the three search tasks combined was $p = .005$. The negative correlation means that, in general, the subjects who received the mental demand manipulation tended to make fewer queries (2.07 on average) than subjects who did not receive the mental demand manipulation (4.04 on average).

Table 4.10

Statistics of Query Reformulation by Each Manipulation and t-test for Query Reformulation by Mental Demand, Temporal Demand, and Frustration for All Three Tasks in Total.

<table>
<tr><th rowspan="2"></th><th rowspan="2">GP</th><th rowspan="2">N</th><th rowspan="2">M</th><th rowspan="2">MD</th><th rowspan="2">SD</th><th rowspan="2">SEM</th><th rowspan="2">SED</th><th rowspan="2">t</th><th rowspan="2">df</th><th rowspan="2">Sig. (2-tailed)</th><th colspan="2">95% CI</th></tr>
<tr><th>LB</th><th>UB</th></tr>
<tr><td rowspan="2">QMD</td><td>CG</td><td>27</td><td>4.04</td><td rowspan="2">1.96</td><td>2.59410</td><td>.49923</td><td rowspan="2">.66754</td><td rowspan="2">2.941</td><td rowspan="2">52</td><td rowspan="2">.005</td><td rowspan="2">.62345</td><td rowspan="2">3.30247</td></tr>
<tr><td>EG</td><td>27</td><td>2.07</td><td>2.30261</td><td>.44314</td></tr>
<tr><td rowspan="2">QTD</td><td>CG</td><td>27</td><td>3.85</td><td rowspan="2">1.70</td><td>3.10958</td><td>.59844</td><td rowspan="2">.73846</td><td rowspan="2">2.307</td><td rowspan="2">52</td><td rowspan="2">.025</td><td rowspan="2">.22188</td><td rowspan="2">3.18553</td></tr>
<tr><td>EG</td><td>27</td><td>2.15</td><td>2.24814</td><td>.43265</td></tr>
<tr><td rowspan="2">QFR</td><td>CG</td><td>27</td><td>5.48</td><td rowspan="2">3.37</td><td>3.52322</td><td>.67804</td><td rowspan="2">.75051</td><td rowspan="2">4.491</td><td rowspan="2">52</td><td rowspan="2">.000</td><td rowspan="2">1.86437</td><td rowspan="2">4.87637</td></tr>
<tr><td>EG</td><td>27</td><td>2.11</td><td>1.67179</td><td>.32174</td></tr>
</table>

Note. M = Mean; MD = Mean Difference; SD = Std. Deviation; SEM = Std. Error Mean; SED = Std. Error Difference, QMD = Query Reformulations by Mental Demand; QTD = Query Reformulations by Temporal Demand; and QFR = Query reformulations by Frustration, GP = Group.

Sub-research question 2 (Temporal Demand Manipulation)

Is there any difference in the propensity for query reformulation behavior when temporal demand increases?

For this question, a t-test was employed (Table 4.10 above). There was a statistically significant negative correlation between temporal demand manipulation and the propensity for query reformulation ($p = .025$). The subjects who received the temporal demand manipulation tended to make fewer queries than the subjects who did not receive temporal demand manipulation.

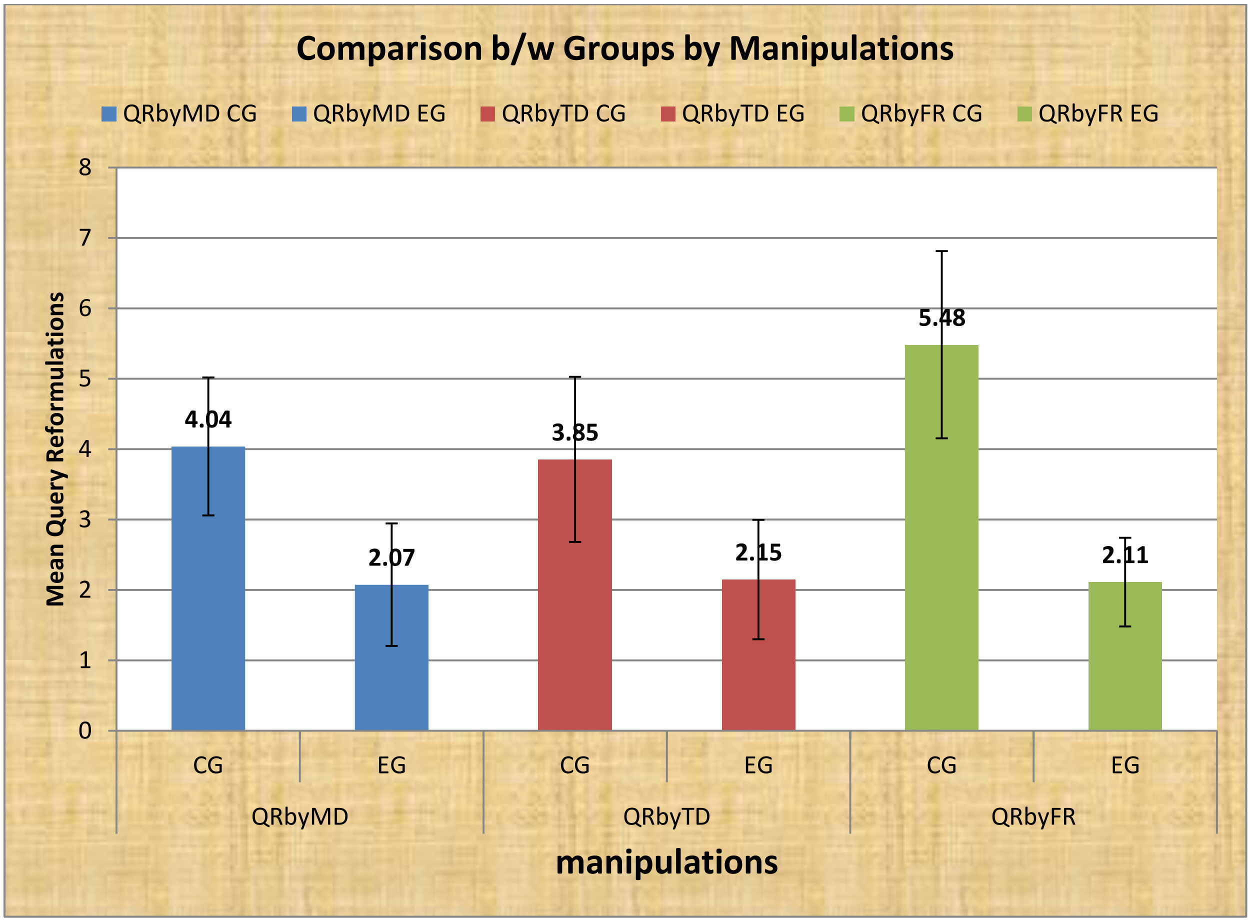

Figure 4.7. Comparison between groups with and without manipulations. Error bars show the 95% confidence intervals for the mean.

Sub-research question 3 (Frustration Manipulation)

Is there any difference in the propensity for query reformulation behavior when frustration increases?

For this question a t-test was employed. There was a statistically significant difference between frustration manipulation and the propensity for query reformulation (p = .000) (Table 4.10). The subjects who received the frustration manipulation tended to make fewer queries than the subjects who did not receive the frustration manipulation.

The results of this study support the findings of Gwizdka (2010). Searchers who experience high cognitive load manipulated by mental demand, temporal demand, or frustration generate fewer queries. The scores on the NASA-TLX also demonstrate that higher cognitive load has negative relationship with the propensity for query reformulation. Therefore, the propensity for query reformulations clearly has a relationship with cognitive load.

NASA-TLX Workload Assessment

The NASA-TLX (Task Load Index) workload assessment was employed to measure how the subjects experienced cognitive load during the experiment. After each search task, subjects completed subjective workload assessments. To this end, the study used the NASA-TLX, in which the subject provided subjective ratings of the workload in six components that contribute to cognitive load (i.e., mental demand, physical demand, temporal demand, performance, effort, and frustration). Each subscale is a seven-point Likert scale, ranging from Extremely Low (1) to Extremely High (7) and the subjects assessed the importance of these components. It asked the subjects the questions outlined in Table 4.11 below.

Table 4.11

NASA-TLX Rating Scale and Description

Components	Scale	Description
Mental demand	extremely low/ extremely high	Rate your level of mental and perceptual activity (i.e., thinking, looking, searching, or remembering) while performing the search task. How mentally demanding was it?
Physical demand	extremely low/ extremely high	Rate your level of physical activity (i.e., clicking, scrolling, or typing) while performing the search task. How physically demanding was it?
Temporal demand	extremely low/ extremely high	Rate the level of perceived time pressure to complete the search task? How hurried or rushed was the pace of the task? How much pressure did you feel to complete the task quickly?
Performance	extremely low/ extremely high	Rate your level of performance in performing the search task. How successful were you in accomplishing what you were asked to do? How satisfied were you with your performance?

Table 4.11 - continued

Components	Scale	Description
Effort	extremely low/ extremely high	Rate your level of effort as you performed the task. How hard did you have to work (mentally and physically) to accomplish it?
Frustration	extremely low/ extremely high	Rate your level of frustration as you performed the task. How insecure, discouraged, irritated, stressed, and annoyed did you feel?

Table 4.12

Summary for Six Components of NASA-TLX, Means, Standard Deviation, Stand Error, 95% Confidence, and Minimum and Maximum

		N	Mean	Std. Deviation	Std. Error	95% CI for Mean		Min	Max
						LB	UB		
MWL_ave	CG	27	4.0123	.80851	.15560	3.6925	4.3322	2.00	5.00
	EG	27	4.7407	1.09128	.21002	4.3090	5.1724	1.00	6.33
TWL_ave	CG	27	3.0864	1.08049	.20794	2.6590	3.5138	1.00	5.00
	EG	27	4.3210	1.16751	.22469	3.8591	4.7828	1.67	6.00
FWL_ave	CG	27	2.8519	.93978	.18086	2.4801	3.2236	1.00	4.67
	EG	27	4.0988	1.35815	.26138	3.5615	4.6360	1.00	6.33
PER_ave	CG	27	4.6667	.88675	.17066	4.3159	5.0175	2.67	6.33
	EG	27	3.7037	1.14479	.22031	3.2508	4.1566	1.67	6.33
EFF_ave	CG	27	4.0741	.64935	.12497	3.8172	4.3309	2.00	5.00
	EG	27	4.8642	1.12569	.21664	4.4189	5.3095	1.00	6.33
PHY_ave	CG	27	3.0000	1.06217	.20441	2.5798	3.4202	1.00	4.33
	EG	27	3.7037	1.15593	.22246	3.2464	4.1610	1.00	5.67

Note. MWL_ave = Mental Demand Workload on average, TWL_ave = Temporal Demand Workload on average, FWL_ave = Frustration Workload on average, PER_ave = Performance on average, EFF_ave = Effort on average, and PHY_ave = Physical Demand Workload on average; CG = Control Group, EG = Experimental Group. *Note*. [a]n = 27; CI = confidence interval; *LB* = Lower Bound, *UB* = Upper Bound; MIN = Minimum, MAX = MAXIUM.

On average, the subjects in experimental group rated their mental demand at M = 4.74 with SD = 1.09128, and SD = .21002, while the subjects in control group rated their mental demand an M= 4.01 with SD = .80851, and SD = .155560. This illustrates that the subjects in the experimental group experienced more mental demand than the subjects in control group (See Figure 4.9).

Among the six components of NASA-TLX, the level of mental workload and effort in experimental group seemed to be dramatically higher than that for control group, but the other components, were rated higher by the control group. The level of performance in the experimental group was lower than that in the control group, which shows that increased cognitive load interferes with performance level (See Table 4.12 & Figure 4.8).

As discussed earlier, subjects' propensity for query reformulation behavior differed significantly between the experimental and control groups, and subjects' NASA-TLX scores differed significantly between the experimental and control groups. The NASA-TLX questionnaire accurately reflects workload changes between groups and identifies individuals who are more likely to experience high workload, and who are more prone to experience poor performance and dissatisfaction during information seeking.

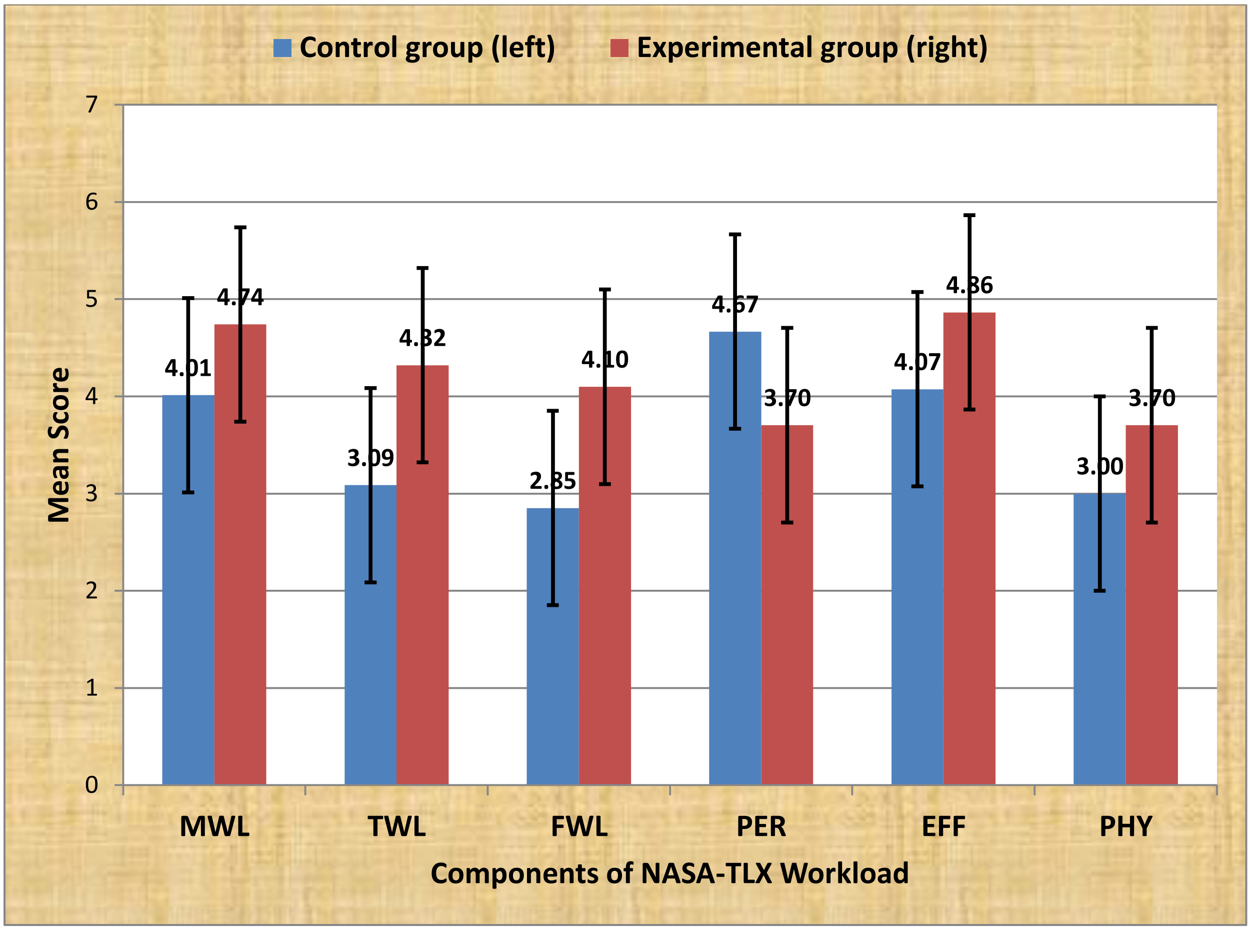

Figure 4.8. Mean score of each component of NASA-TLX workload. MWL = Mental Demand Workload; TWL = Temporal Demand Workload; FWL = Frustration Workload; PER = Performance; EFF = Effort; PHY = Physical Demand Workload. Error bars show the 95% confidence intervals for the mean.

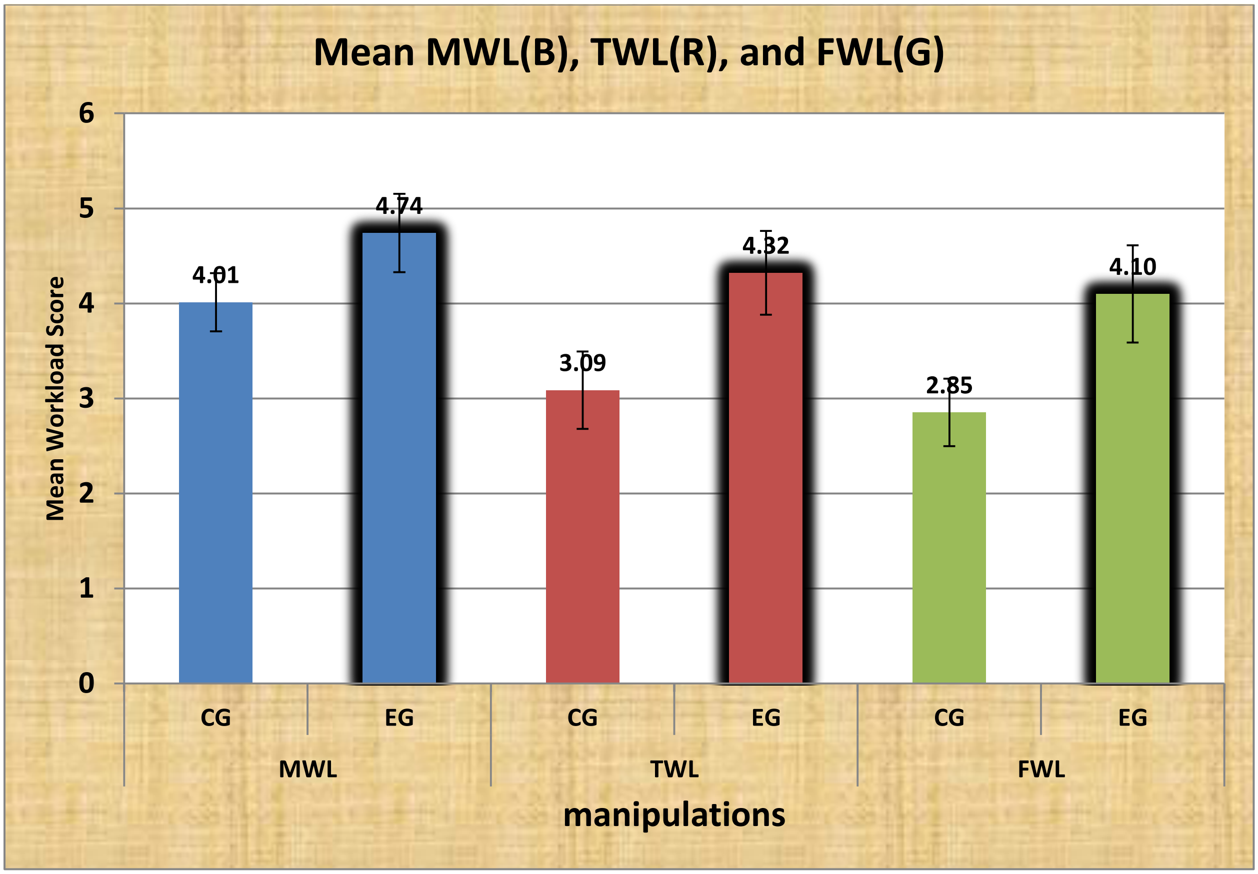

Figure 4.9. Mean workload score by manipulations. MWL = Mental Workload Score, TWL = Temporal Workload Score, and FWL = Frustration Score. CG = Control Group; EG = Experimental Group. Error bars show the 95% confidence intervals for the mean.

Total Score of the NASA-TLX

Cognitive load may vary due to each components of NASA-TLX workload assessment (i.e., mental demand, temporal demand, physical demand, performance, effort, and frustration). The total score of NASA-TLX in this study was calculated as follows:

Total NASA-TLX score = MWL + TWL + FWL + PER + EFF + PHY, where MWL = mental work load, TWL = temporal workload, FWL = frustration, PER = performance, EFF = effort, and PHY = physical workload.

The mean value of the total score for the experimental group was M = 25.43 with SD = 4.34005 and SE = .83524, while the mean value for the control group was M = 21.69 with SD = 2.88220 and SD = .55468.

Table 4.13

Summary for Mean of Total NASA_TLX Workload Assessment, Means, Standard Deviation, Stand Error, 95% Confidence, and Minimum and Maximum

Group	N	Mean	Std. Deviation	Std. Error	95% CI for Mean		Min	Max
					Lower Bound	Upper Bound		
CG	27	21.69	2.88220	.55468	20.5512	22.8315	14.33	26.00
EG	27	25.43	4.34005	.83524	23.7152	27.1490	12.00	32.00

Note. CG = Control Group; EG = Experimental Group, CI = Confidence Interval; Min = Minimum; Max = Maximum.

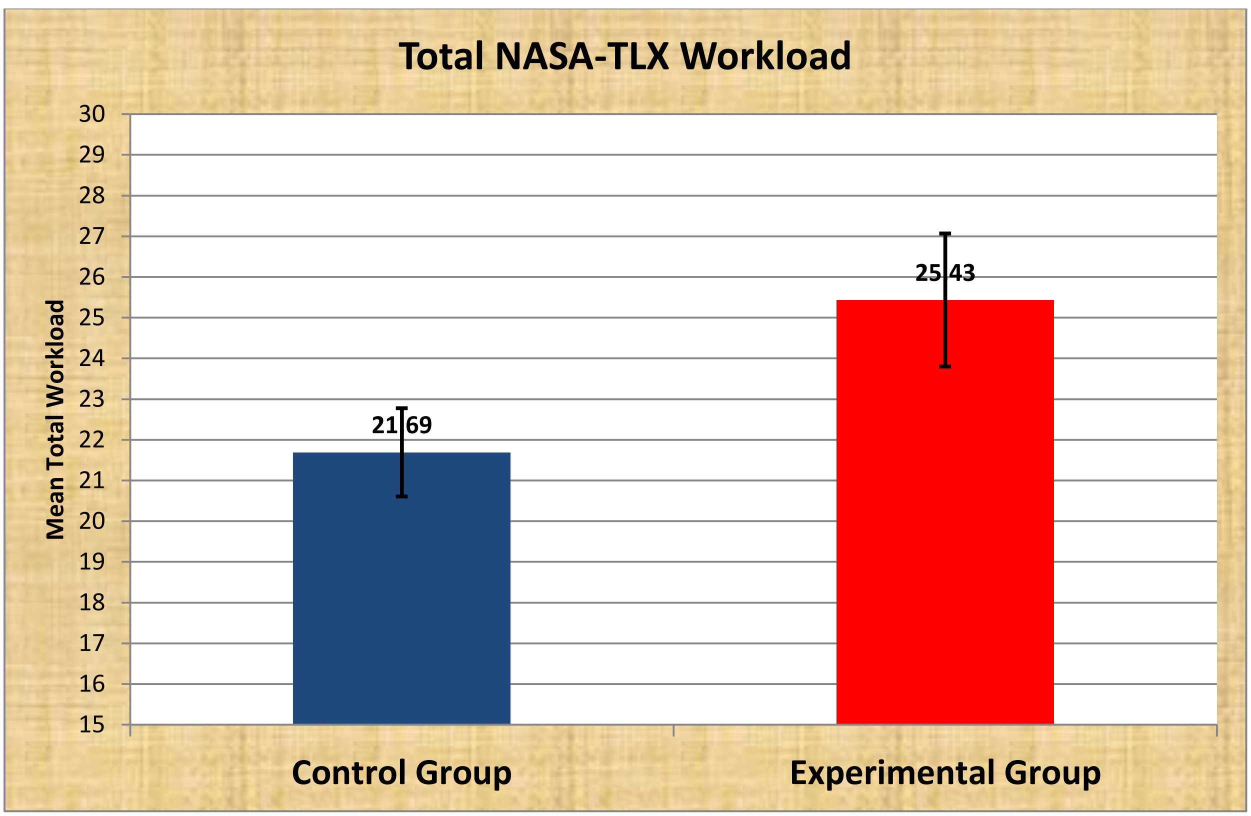

Figure 4.10. Mean total NASA-TLX workload assessment by groups. Error bars show the 95% confidence intervals for the mean.

Table 4.14

Summary for ANOVA, Total NASA_TLX Workload Assessment between Groups

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	188.907	1	188.907	13.92	**.000**
Within Groups	705.720	52	13.572		
Total	894.628	53			

Note. p <.05 in 95% Confidence Interval.

As seen in Table 15 and Figure 4.10, the ANOVA results show that there is a significant difference $F(1, 52) = 13.92$ in $p < 0.05$ between groups in terms of the total score of NASA-TLX workload assessment, which means that the experimental group experienced more cognitive load than the control group.

Table 4.15

Summary for ANOVA for Mean of Each Component of NASA-TLX Workload

		Sum of Squares	df	Mean Square	F	Sig.
MWL	Between Groups	7.163	1	7.163	7.766	**.007**
	Within Groups	47.959	52	.922		
	Total	55.121	53			
TWL	Between Groups	20.576	1	20.576	16.262	**.000**
	Within Groups	65.794	52	1.265		
	Total	86.370	53			
FWL	Between Groups	20.990	1	20.990	15.390	**.000**
	Within Groups	70.922	52	1.364		
	Total	91.912	53			
PER	Between Groups	12.519	1	12.519	11.940	**.001**
	Within Groups	54.519	52	1.048		
	Total	67.037	53			
EFF	Between Groups	8.428	1	8.428	9.981	**.003**

Table 4.15 - continued

		Sum of Squares	df	Mean Square	F	Sig.
	Within Groups	43.909	52	.844		
	Total	52.337	53			
PHY	Between Groups	6.685	1	6.685	5.425	**.024**
	Within Groups	64.074	52	1.232		
	Total	70.759	53			

Note. MWL = Mental Workload; TWL = Temporal Workload; FWL = Frustration; PER = Performance; EFF = Effort; PHY = Physical Workload, 95% Confidence Interval.

As shown in Table 4.15 above, each component of the NASA-TLX cognitive load scores of searchers who were exposed to the three cognitive load manipulations was significantly greater than those of searchers who were not exposed (i.e., p = .007 in MWL, p = .000 in TWL, p = .000 in FWL, p = .001 in PER, p = 003 in EFF, and p = .024 in PHY). The ANOVA in Table 4.16 shows that this test is significant with $p < .05$. Given that our model represents the group differences, this ANOVA tells us that using group means to predict scores is significantly better than using the overall mean, i.e. the group means are significantly different.

Manipulations

This study focused on the manipulation of three components that contribute to cognitive load: mental demand, temporal demand, and frustration. Among these three components, temporal demand manipulation showed the highest number of query reformulations (M=2.15, SE=.433), followed by frustration (M=2.11, SE=.322), and mental demand (M=2.07, SE=.443) (See Table 4.16).

The number of query reformulations made by the experimental group when receiving the frustration manipulation was 2.6 times less than that of the control group. The number of query reformulations made by the experimental group when receiving the mental demand manipulation was 1.97 times less compared to the control group. The number of query reformulations made by the experimental group when receiving the temporal demand manipulation experimental group queries was 1.79 times less compared to the control group. Among the three manipulations, frustration decreased the number of query reformulations the most.

Table 4.16

Mean Query Reformulations for All Three Search Tasks in Total

factor	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
QRbyMD	2.07	.443	1.163	2.985
QRbyTD	2.15	.433	1.259	3.037
QRbyFR	2.11	.322	1.450	2.772

Note. MD = Mental demand, TD = Temporal demand, and FR = Frustration; QR = Query Reformulation.

Mental demand

As suggested in previous studies, the mental demand manipulation succeeded at increasing the mental demand of the experimental group in this study. Increasing the mental demand in the experimental group generated fewer query reformulations than those generated without a mental demand manipulation (control group). Bruken, Plass, and Leutner (2003) state that "although no single measure is ideal, we believe that adding a dual-task approach to the existing measures of cognitive load will allow for a more valid and reliable assessment of cognitive load" (p.60). Recarte and Nunes (2003) examined the effects of mental workload on visual search and decision making in real traffic conditions with 12 participants who drove a car with an eye-tracking system. Mental workload was manipulated by having participants perform several mental tasks while driving. They revealed that mental tasks produced spatial gaze concentration and visual-detection impairment. The findings in this study are consistent with this previous research.

Temporal demand

The temporal demand manipulation worked well at increasing the level of temporal demand experienced by the experimental group. Block, Hancock, and Jakay (2010) state that "the types of cognitive load manipulations that moderated prospective duration judgment are those that clearly demand working-memory resources, especially those that involve a hypothetical central executive" (p. 10). Limiting the experimental group's time generated fewer query reformulations than generated by the control group. This is consistent with previous work (Birch, Juul-Kristensen, Jensen, Finsen, & Christensen, 2000) studying the effect of time pressure and mental demand on acute response during computer drawing work. Birch, et al. (2000) found that time pressure combined with low precision and low mental demand resulted in higher eletromyographic (EMG) activity for all muscles and in a small increase in the number of drawings produced.

Frustration

The frustration manipulation worked properly by impeding query reformulations while searchers performed a search task. According to Pew Internet and American Life Project (2011), the use of information technology comes frustration with advantages, 20% of cell owners experienced frustration because their phone was taking too long to download something; 16% had difficulty reading something on their phone because the screen was too small; and 10% had difficulty entering a lot of text on their phone. Jansen (1998) states that "Bad GUI [Graphic User Interface] design prevents the user from concentrating on the primary cognitive task. This results in user frustrations, decreased performance, higher costs, and possibly product and marketplace failure" (p. 26). There are different types of mobile computing devices, such as PDAs (personal digital assistants), tablet personal computers, and Smartphones that provide many different input modalities for information access and retrieval. However, as mobile technologies have evolved and become smaller, their limitations in terms of input and output (small visual display, use of fingers to operate buttons, and the lack of an alphanumeric keyboard screen and mouse) become apparent (Gu & Gilbert, 2004). Baecker, Booth, Jovicic, McGrenere, and Moore (2000) explain that poorly crafted interfaces and system complexity cause negative emotional responses, which lead to experiences of frustration, confusion, and failure. A user may become frustrated when having trouble finding information with a search engine or information system even though the user is ultimately successful. Searchers' emotions can be affected by search frustration and usability frustration. Such experiences may be felt most strongly by users who are confronted with the inability to find information in a timely manner, and with difficulty in using the features provided by a search engine or information system that they cannot easily understand. Field, Allan, and Jones (2010) state that by modeling searcher frustration, search engines can predict the current state of user frstration and decide when to intervene with alternative search strategies to prevent the user from becoming more frustrated, giving up, or switching to another search engine. The results of this study show that searchers made fewer queries when using an on-screen keyboard during information search process.[2] The physical demand of information search activities should not be overlooked. Declining cognitive and physical functioning may impede performance of query reformulation and search results. This may indicate that subjects want to use a system that is easy to use and has simple usability, functionality, and accessibility.

Validation Check for Manipulations

To ensure that the three manipulations of cognitive load worked properly, a validation step was performed. It is necessary to verify that each treatment of the three manipulations actually had two distinguishable levels of difficulty—a high and low score. Repeated measures outcomes in the experimental group were examined to ensure that each manipulation worked as expected. The following questions were checked with repeated measures:

1. When mental demand is manipulated, is the mental demand score higher than the other two manipulations?
2. When temporal demand is manipulated, is the temporal demand score higher than the other two manipulations?
3. When frustration is manipulated, is the frustration score higher than the other two manipulations?

As shown in Table 4.17 below, the mental demand score in the mental demand manipulation (M = 5.33, SD = 1.49) is higher than the mental demand score in the temporal demand manipulation (M = 4.81, SD = 1.30) and the frustration manipulation (M = 4.26, SD = 1.50). Similarly, the temporal demand score in the temporal demand manipulation (M = 5.41, SD = 1.28) is higher

[2] It might be interesting to examine how mobile application users can differ in terms of the propensity of query reformulations since it has smaller size of input device than that of desktop computer or notebook computer.

than the temporal demand in the mental demand manipulation (*M* = 4.04, *SD* = 1.58) and the frustration manipulation (*M* = 3.52, *SD* = 1.53). However, the frustration score in the frustration manipulation (*M* = 3.93, *SD* = 1.90) is not higher than the frustration score in the mental demand manipulation (*M* = 4.07, *SD* = 1.77) and the temporal demand manipulation (*M* = 4.30, *SD* = 1.68). It is interesting to see that the propensity for query reformulation in the frustration manipulation is less than with the other two manipulations.

In light of the results of the pilot study and the test, it is likely that the manipulations of mental demand and temporal demand were successful and valid, but the manipulation of frustration was only partially successful. Subjects' ratings of the mental workload, temporal workload, and frustration workload were tested to validate the manipulation of cognitive load in each search task (See Table 4.17 and Table 4.19).

Among three manipulations, the temporal demand score (5.41) is higher than the scores for the other two manipulations: mental demand (5.33) and frustration (3.93) (See Table 4.17 and Figure 4.11).

Table 4.17

Summary of Mean Workload Manipulations of Experimental Group for Validation Check

	Mean	Std. Deviation	N
Mental workload in mental demand condition	**5.33**	1.49	27
Mental workload in temporal demand condition	4.81	1.30	27
Mental workload in frustration condition	4.23	1.51	27
Temporal workload in mental demand condition	4.04	1.58	27
Temporal workload in temporal demand condition	**5.41**	1.28	27
Temporal workload in frustration condition	3.52	1.53	27
Frustration workload in mental demand condition	4.07	1.77	27
Frustration workload in temporal demand condition	4.30	1.68	27
Frustration workload in frustration condition	**3.93**	1.90	27

Note. 0 = Extremely low, 7 = Extremely high.

Table 4.18

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	4720.333	1	4720.333	560.689	**.000**	.956	560.689	1.000
Error	218.889	26	8.419					
a. Computed using alpha = .05								

Table 4.19

Correlation for Type of Rating and Type of Workload

Typerating	Typewkload	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
MWL	by MD manipulation	5.33	.287	4.742	5.924
	by TD manipulation	4.82	.251	4.300	5.330
	by FR manipulation	4.26	.290	3.662	4.856
TWL	by MD manipulation	4.04	.304	3.412	4.662
	by TD manipulation	5.41	.246	4.902	5.913
	by FR manipulation	3.52	.294	2.914	4.123
FRW	by MD manipulation	4.07	.341	3.372	4.776
	by TD manipulation	4.30	.324	3.631	4.962
	by FR manipulation	3.93	.366	3.174	4.678

Note. Typerating = NASA-TLX score of mental demand, temporal demand, and frustration; Typewkload = mental demand, temporal demand, and frustration. MD = Mental Demand; TD = Temporal Demand; FR = Frustration, 0 = Extremely low, 7 = Extremely high.

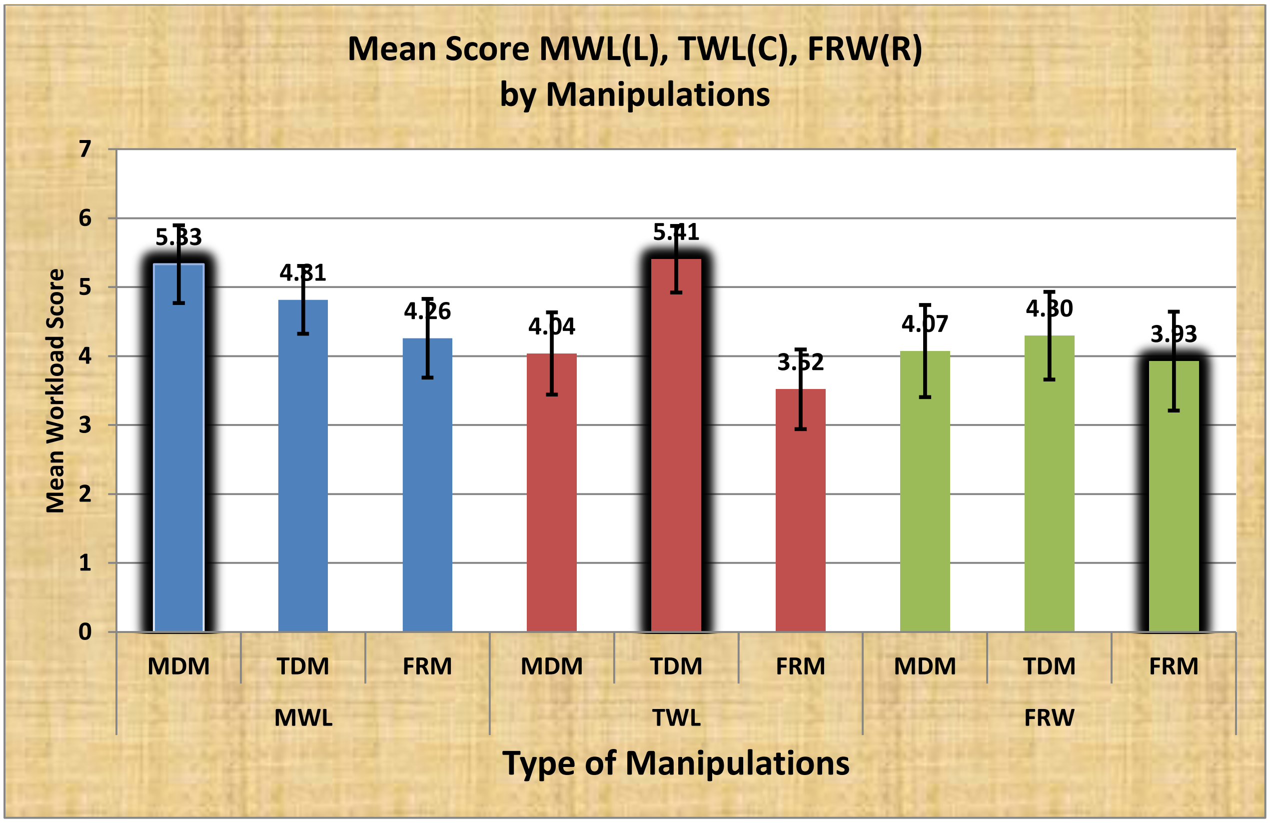

Figure 4.11. Mental Demand, Temporal Demand, and Frustration Workload score depending upon three different manipulations. 1. MWL = Mental Demand Workload Score; 2.TWL = Temporal Demand Workload Score; 3. FRW = Frustration Score. MDM = mental demand manipulation; TDM = Temporal demand manipulation; FRM = Frustration Manipulation, 0 = Extremely low, 7 = Extremely high. Error bars show the 95% confidence intervals for the mean.

Table 4.20

Pairwise Comparison by type of workload (MD, TD, and FR)

(I) typewkload	(J) typewkload	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
MD	TD	.481	.202	.075	-.036	.999
	FR	.704*	.205	**.006**	.180	1.227
TD	MD	-.481	.202	.075	-.999	.036
	FR	.222	.218	.952	-.335	.780
FR	MD	-.704*	.205	**.006**	-1.227	-.180
	TD	-.222	.218	.952	-.780	.335

Note. MD = Mental Demand; TD = Temporal Demand; FR = Frustration.

As shown in table 4.20 above and Figure 4.11, mental demand is associated with frustration. Future studies should examine this relationship more closely.

The results of this study show that the propensity for query reformulations is affected by cognitive load. Conditions of high cognitive load generated fewer queries, while conditions of low cognitive load generated more queries. However, it does not necessarily follow that more queries are always better than fewer queries in terms of performance; rather queries may increase the number of interactions between the system and user.

Overall Query Reformulation Experience

After each search task was completed, subjects were asked to complete a post-task survey questionnaire that included overall experience with query reformulation during the experiment. The questions are outlined in Table 4.21. Chi-Square tests were conducted of overall query reformulation experience for each task and show that there are no significant differences between the control group and the experimental group in task1, task2, and task3 in terms of the questions: when do you think you needed to reformulate your queries?, why do you think you reformulated your previous queries?, and in what way did the query reformulation affect the search results throughout a search? However, there is a difference between groups in terms of when they thought they needed to reformulate their queries in task2 with Pearson Chi-Square value 10.113 at p = .039 (See Table 4.33).

Overall, except for one question related to the library search task (i.e., when do you think you needed to reformulate your queries?), there are no statistically significant differences in the responses of the experimental group and the control group.

Table 4.21

Overall Query Reformulation Experience

Variables	Option	Description
When	When there were too many search results	When do you think you needed to reformulate your queries?
	When there were too few search results	
	When the search results are not satisfactory	
	To better find relevant information	
	I don’t know	
Why	To enhance relevance	Why do you think you reformulated your previous queries?
	To enhance recall	
	To enhance precision	
	To make a comparison between the search results of queries	
	To feel more satisfied and confident	
In what ways	It improved the search results	In what ways did the query reformulations affect the search results throughout a search?
	It narrowed the search results	
	It broadened the search results	
	It inspired new thinking/ ideas	
	It wasn’t helpful; it wasted my time	

Task1 (Virus search task) query reformulation experience

Detailed information about the virus search task query reformulation experience is provided below.

Table 4.22

Summary of Descriptive Statistics of Chi-Square on query reformulation experience in terms of when QR needed in task1 in between Control vs. Experimental group

		when to QR needed in task1 1=too many, 2=too few, 3=not satisfactory, 4=better find relevant info, 5=I don't know					
		too many search results	too few search results	search results are not satisfactory	to better find relevant information	i don't know	Total
control vs experimental	0	7	0	9	9	1	26
	1	3	1	10	11	0	25
Total		10	1	19	20	1	51

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There are 3missing cases out of 54.

When asked "when do you think you needed to reformulate your previous queries" after completing task1 (i.e., virus search task), 9 subjects in control group and 11 subjects in experimental group answered that they reformulated previous queries because they wanted to better find relevant information. Nine subjects in control group and 10 subjects in experimental group answered that they reformulated previous queries because search results were not satisfactory. Seven subjects in control group and 3 subjects in experimental group answered that they reformulated previous queries because there were too many search results. None of subjects in control group but only one subject in experimental group answered that they reformulated previous queries because there were too few search results.

This table 4.22 shows us that there is no statistically significant association between group and preferred reason to reformulating previous queries. However, relevance seems to be the key factor for subjects when considering reformulating previous queries followed by search satisfaction. 10 subjects (i.e., 7 in control group and 3in experimental group) needed query reformulations when there were too many search results.

Table 4.23

Summary of Chi-Square on query reformulation experience in terms of when QR needed in task1 in between Control vs. Experimental group

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.834[a]	4	.429
Likelihood Ratio	4.652	4	.325
Linear-by-Linear Association	.744	1	.388
N of Valid Cases	51		
a. 5 cells (50.0%) have expected count less than 5. The minimum expected count is .49.			

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There are 3missing cases out of 54.

We can also see here in table 4.23 that Chi-square value = 3.834, p = 0.429. This tells us that there is no statistically significant association between group (i.e., control vs. experimental) and the preferred reason of when to QR needed (i.e., 1= too many search results, 2= too few search results, 3=search results are not satisfactory, 4=to better find relevant information, 5=I don't know) in task1 (i.e., virus search task).

Table 4.24

Summary of Descriptive Statistics of Chi-Square on query reformulation experience in terms of why QR in task1 in between Control vs. Experimental group

		why QR in task1 1=increase relevance, 2=increase recall, 3=increase precision, 4=make comparisons, 5=feel more satisfied					
		to increase relevance	to increase recall	to increase precision	to make a comparison between search results	to feel more satisfied and confident	Total
control vs experimental	0	6	0	10	7	3	26
	1	8	1	10	4	2	25
Total		14	1	20	11	5	51

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There are 3cases missing out of 54.

When asked "why do you reformulated your previous queries" after completing task1 (i.e., virus search task), ten subjects in control group and 10 subjects in experimental group answered that they reformulated previous queries because they wanted to increase precision. Six subjects in control group and 8 subjects in experimental group answered that they reformulated previous queries because they wanted to increase relevance. Seven subjects in control group and 4 subjects in experimental group answered that they reformulated previous queries because they wanted to make a comparison between search results. Three subjects in control group and 2 subjects in experimental group answered that they reformulated previous queries because they wanted to feel more satisfied and confident. None of subjects in control group but only one subject in experimental group answered that they reformulated previous queries because he/she wanted to increase recall.

This table 4.24 shows us that there is no statistically significant association between group and preferred reason to reformulating previous queries. However, precision and relevance seem to be two key factors for subjects to reformulate previous queries as well as to compare between search results. Interestingly, search satisfaction and confidence might be other factors for subjects when considering query reformulations.

Table 4.25

Summary of Chi-Square on query reformulation experience in terms of why QR in task1 in between Control vs. Experimental group

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.285[a]	4	.683
Likelihood Ratio	2.683	4	.612
Linear-by-Linear Association	1.166	1	.280
N of Valid Cases	51		
a. 4 cells (40.0%) have expected count less than 5. The minimum expected count is .49.			

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There are 3cases missing out of 54.

We can see here in table 4.25 that Chi-square value = 2.285, p = 0.683. This tells us that there is no statistically significant association between group (i.e., control vs. experimental) and the preferred reason of why QR in task1 (i.e., 1=increase relevance, 2=increase recall, 3=increase precision, 4=make comparisons, 5=feel more satisfied) in task1 (i.e., virus search task).

Table 4.26

Summary of Descriptive Statistics of Chi-Square on query reformulation experience in terms of in what ways QR affect the search results in task1 in between Control vs. Experimental group

		in what ways QR affects in task1 1=improved search results, 2= narrowed search results, 3=broadened search results, 4=inspired new ideas, 5=wasn't helpful					
		improved the search results	narrowed the search results	broadened the search results	inspired new thinking or ideas	wasn't helpful; it wasted my time	Total
control vs experimental	0	9	6	2	5	4	26
	1	9	10	2	4	0	25
Total		18	16	4	9	4	51

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There are 3cases missing out of 54.

When asked "in what ways did the query reformulations affect the search results throughout a search" after completing task1 (i.e., virus search task), 9 subjects in control group and 9 subjects in experimental group answered that query reformulations improved the search results. Six subjects in control group and 10 subjects in experimental group answered that query reformulations narrowed the search results. None of subjects in experimental group but only 4 subjects in control group answered that query reformulations were not helpful and waste of their time.

This table 4.26 shows us that there is no statistically significant association between group and preferred reason to reformulating previous queries. However, precision and relevance seem to be two key factors for subjects to reformulate previous queries. Interestingly, query reformulations could inspire new thinking or ideas.

Table 4.27

Summary of Chi-Square on query reformulation experience in terms of in what ways QR affect the search results in task1 in between Control vs. Experimental group

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5.093[a]	4	.278
Likelihood Ratio	6.648	4	.156
Linear-by-Linear Association	2.065	1	.151
N of Valid Cases	51		
a. 6 cells (60.0%) have expected count less than 5. The minimum expected count is 1.96.			

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There are 3cases missing out of 54

The value of the chi-square statistic is given in the table 4.27 (and the degrees of freedom) as is the significance value. The value of the chi-square statistic is 5.093. This value is not highly significant ($p < .001$), indicating that the type of group did not have a significant effect on in what ways query reformulations would affect the search results.

Task2 (Library search task) query reformulation experience

Detailed information about library search task query reformulation experience is provided below.

Table 4.28

Summary of Descriptive Statistics of Chi-Square on query reformulation experience in terms of when QR needed in task2 in between Control vs. Experimental group

		when to QR needed in task2 1=too many, 2=too few, 3=not satisfactory, 4=better find relevant info, 5=I don't know				
		1 too many search results	3 search results are not satisfactory	4 to better find relevant information	5 i don't know	Total
control vs experimental	0	6	10	9	2	27
	1	1	9	13	3	26
Total		7	19	22	5	53

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1missing case out of 54.

When asked "when do you think you needed to reformulate your previous queries" after completing task2 (i.e., library search task), 9 subjects in control group and 13 subjects in experimental group answered that they reformulated previous queries because they wanted to better find relevant information. Ten subjects in control group and 9 subjects in experimental group answered that they reformulated previous queries because search results were not satisfactory. Six subjects in control group and only one subjects in experimental group answered that they reformulated previous queries because there were too many search results. None of subjects in both group answered that they reformulated previous queries because there were too few search results.

This table 4.28 shows us that there is no statistically significant association between group and preferred reason to reformulating previous queries. However, relevance seems to be the key factor for subjects when considering reformulating previous queries followed by search satisfaction.

Table 4.29

Summary of Chi-Square on query reformulation experience in terms of when QR needed in task2 in between Control vs. Experimental group

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4.534[a]	3	.209
Likelihood Ratio	4.929	3	.177
Linear-by-Linear Association	4.102	1	.043
N of Valid Cases	53		
a. 4 cells (50.0%) have expected count less than 5. The minimum expected count is 2.45.			

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1missing case out of 54.

We can also see here in table 4.29 that Chi-square value is 4.534 at p = 0.209. This tells us that there is no statistically significant association between group (i.e., control vs. experimental) and the preferred reason to reformulating previous queries needed (i.e., 1= too many search results, 2= too few search results, 3=search results are not satisfactory, 4=to better find relevant information, 5=I don't know) in task2 (i.e., library search task).

Table 4.30

Summary of Descriptive Statistics of Chi-Square on query reformulation experience in terms of why QR in task2 in between Control vs. Experimental group

		why QR in task2 1=increase relevance, 2=increase recall, 3=increase precision, 4=make comparisons, 5=feel more satisfied				
		1 to increase relevance	3 to increase precision	4 to make a comparison between search results	5 to feel more satisfied and confident	Total
control vs experimental	0	8	7	4	8	27
	1	14	5	3	4	26
Total		22	12	7	12	53

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1 case missing out of 54.

When asked "why do you reformulated your previous queries" after completing task2 (i.e., library search task), 8 subjects in control group and 14 subjects in experimental group answered that they reformulated previous queries because they wanted to increase relevance. Seven subjects in control group and 5 subjects in experimental group answered that they reformulated previous queries because they wanted to increase precision. Four subjects in control group and 3 subjects in experimental group answered that they reformulated previous queries because they wanted to make a comparison between search results. Eight subjects in control group and 4 subjects in experimental group answered that they reformulated previous queries because they wanted to feel more satisfied and confident. None of subjects in control group but only one subject in experimental group answered that they reformulated previous queries because he/she wanted to increase recall. Relevance and precision seem to be two key factors for subjects to reformulate previous queries as well as comparing between search results. Interestingly, search satisfaction and confidence might be other factors for subjects why they were considering query reformulations.

Table 4.31

Summary of Chi-Square on query reformulation experience in terms of why QR in task2 in between Control vs. Experimental group

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.428[a]	3	.330
Likelihood Ratio	3.476	3	.324
Linear-by-Linear Association	3.169	1	.075
N of Valid Cases	53		
a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 3.43.			

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1case missing out of 54.

We can see here in table 4.31 that Chi-square value = 3.428, p = 0.330, indicating that there is no statistically significant association between group (i.e., control vs. experimental) and the preferred reason to reformulating previous queries in task2 (i.e., 1=increase relevance, 2=increase recall, 3=increase precision, 4=make comparisons, 5=feel more satisfied) in virus search task.

Table 4.32

Summary of Descriptive Statistics of Chi-Square on query reformulation experience in terms of in what ways QR affect the search results in task2 in between Control vs. Experimental group

		in what ways QR affects in task2 1=improved search results, 2= narrowed search results, 3=broadened search results, 4=inspired new ideas, 5=wasn't helpful					
		1 improved the search results	2 narrowed the search results	3 broadened the search results	4 inspired new thinkings or ideas	5 wasn't helpful; it wasted my time	Total
control vs experimental	0	12	7	2	5	1	27
	1	9	7	3	0	7	26
Total		21	14	5	5	8	53

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1case missing out of 54.

When asked “in what ways did the query reformulations affect the search results throughout a search” after completing task2 (i.e., library search task), 12 subjects in control group and 9 subjects in experimental group answered that query reformulations improved the search results. Seven subjects in control group and 7 subjects in experimental group answered that query reformulations narrowed the search results. One subject in control group and 7 subjects in experimental group answered that query reformulations were not helpful and a waste of their time. Similarly, precision and relevance seem to be two key factors for subjects to reformulate previous queries. Interestingly, query reformulations could inspire new thinkings or ideas and broaden the search results.

Table 4.33

Summary of Chi-Square on query reformulation experience in terms of in what ways QR affect the search results in task2 in between Control vs. Experimental group

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.113[a]	4	.039
Likelihood Ratio	12.606	4	.013
Linear-by-Linear Association	1.336	1	.248
N of Valid Cases	53		
a. 6 cells (60.0%) have expected count less than 5. The minimum expected count is 2.45.			

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1case missing out of 54.

This table 4.33 shows us that there is statistically significant association between group and preferred reason to reformulating previous queries. The value of the chi-square statistic is given in the table 4.33 (and the degrees of freedom) as is the significance value. The value of the chi-square statistic is 10.113. This value is significant ($p < .05$), indicating that the type of group did have a significant effect on in some ways query reformulations would affect the search results. According to Field (2009), if the significance value is small enough (conventionally Sig. must be less than .05) then we reject the hypothesis that the variables are independent and gain confidence in the hypothesis that they are in some way related (p. 697).

Task3 (Security search task) query reformulation experience

Detailed information about the security search task query reformulation experience is provided below.

Table 4.34

Summary of Descriptive Statistics of Chi-Square on query reformulation experience in terms of when QR needed in task3 in between Control vs. Experimental group

		when to QR needed in task3 1=too many, 2=too few, 3=not satisfactory, 4=better find relevant info, 5=I don't know					
		1 too many search results	2 too few search results	3 search results are not satisfactory	4 to better find relevant information	5 i don't know	Total
control vs experimental	0	7	2	7	10	1	27
	1	2	3	8	12	2	27
Total		9	5	15	22	3	54

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1missing case out of 54.

When asked "when do you think you needed to reformulate your previous queries" after completing task3 (i.e., security search task), 10 subjects in control group and 12 subjects in experimental group answered that they reformulated previous queries because they wanted to better find relevant information. Seven subjects in control group and 8 subjects in experimental group answered that they reformulated previous queries because search results were not satisfactory. Seven subjects in control group and 2 subjects in experimental group answered that they reformulated previous queries because there were too many search results. Two subjects in control group and 3 subjects in experimental group answered that they reformulated previous queries because there were too few search results. Relevance seems to be the key factor for subjects when considering reformulating previous queries followed by search results satisfaction.

Table 4.35

Summary of Chi-Square on query reformulation experience in terms of when QR needed in task3 in between Control vs. Experimental group

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.560[a]	4	.469
Likelihood Ratio	3.732	4	.444
Linear-by-Linear Association	2.225	1	.136
N of Valid Cases	54		
a. 6 cells (60.0%) have expected count less than 5. The minimum expected count is 1.50.			

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1missing case out of 54.

This table 4.35 shows us that there is no statistically significant association between group and preferred reason to reformulating previous queries. We can see here in table 4.35 that Chi-square value = 3.560, p = 0.469, indicating that there is no statistically significant association between group (i.e., control vs. experimental) and the preferred reason of when to QR needed (i.e., 1= too many search results, 2= too few search results, 3=search results are not satisfactory, 4=to better find relevant information, 5=I don't know) in library search task.

Table 4.36

Summary of Descriptive Statistics of Chi-Square on query reformulation experience in terms of why QR in task3 in between Control vs. Experimental group

		why QR in task3 1=increase relevance, 2=increase recall, 3=increase precision, 4=make comparisons, 5=feel more satisfied					
		1 to increase relevance	2 to increase recall	3 to increase precision	4 to make a comparison between search results	5 to feel more satisfied and confident	Total
control vs experimental	0	10	0	11	5	1	27
	1	8	2	8	6	3	27
Total		18	2	19	11	4	54

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1case missing out of 54.

When asked “why do you reformulated your previous queries” after completing task3 (i.e, library search task), 11 subjects in control group and 8 subjects in experimental group answered that they reformulated previous queries because they wanted to increase precision. Ten subjects in control group and 8 subjects in experimental group answered that they reformulated previous queries because they wanted to increase relevance. Five subjects in control group and 6 subjects in experimental group answered that they reformulated previous queries because they wanted to make a comparison between search results. One subject in control group and 3 subjects in experimental group answered that they reformulated previous queries because they wanted to feel more satisfied and confident. None of subjects in control group but only 2 subjects in experimental group answered that they reformulated previous queries because he/she wanted to increase recall. Similarly, precision and relevance seem to be two key factors for subjects to reformulate previous queries as well as to compare between search results. Interestingly, search satisfaction and confidence might be other factors for subjects when considering query reformulations.

Table 4.37

Summary of Chi-Square on query reformulation experience in terms of why QR in task3 in between Control vs. Experimental group

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.787[a]	4	.436
Likelihood Ratio	4.608	4	.330
Linear-by-Linear Association	.510	1	.475
N of Valid Cases	54		
a. 4 cells (40.0%) have expected count less than 5. The minimum expected count is 1.00.			

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1; There is 1case missing out of 54.

This table 4.37 shows us that there is no statistically significant association between group and preferred reason why they reformulated previous queries. We can see here in table 4.37 that Chi-square value = 2.285, p = 0.683. This tells us that there is no statistically significant association between group (i.e., control vs. experimental) and the preferred reason of why QR in task3 (i.e., 1=increase relevance, 2=increase recall, 3=increase precision, 4=make comparisons, 5=feel more satisfied) in library search task.

Table 4.38

Summary of Descriptive Statistics of Chi-Square on query reformulation experience in terms of in what ways QR affect the search results in task3 in between Control vs. Experimental group

		in what ways QR affects in task3 1=improved search results, 2= narrowed search results, 3=broadened search results, 4=inspired new ideas, 5=wasn't helpful					
		1 improved the search results	2 narrowed the search results	3 broadened the search results	4 inspired new thinkings or ideas	5 wasn't helpful; it wasted my time	Total
control vs experimental	0	13	5	3	4	2	27
	1	12	2	6	6	1	27
Total		25	7	9	10	3	54

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1.

When asked "in what ways did the query reformulations affect the search results throughout a search" after completing task3 (i.e., security search task), 13 subjects in control group and 12 subjects in experimental group answered that query reformulations improved the search results. Five subjects in control group and 2 subjects in experimental group answered that query reformulations narrowed the search results. Similarly, relevance and precision seem to be two key factors for subjects to reformulate previous queries. Interestingly, query reformulations could inspire new thinkings or ideas and broaden the search results.

Table 4.39

Summary of Chi-Square on query reformulation experience in terms of in what ways QR affect the search results in task3 in between Control vs. Experimental group

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.059[a]	4	.548
Likelihood Ratio	3.130	4	.536

Table 4.39 - continued

Chi-Square Tests			
	Value	df	Asymp. Sig. (2-sided)
Linear-by-Linear Association	.251	1	.617
N of Valid Cases	54		
a. 6 cells (60.0%) have expected count less than 5. The minimum expected count is 1.50.			

Note. Task1 = virus search task; task2 = library search task; task3 = security search task, Control Group = 0, Experimental Group =1.

This table 4.39 shows us that there is no statistically significant association between group and preferred reason to reformulating previous queries. The value of the chi-square statistic is given in the table 4.39 (and the degrees of freedom) as is the significance value. The value of the chi-square statistic is 3.059 at p=0.548. This value is not highly significant indicating that the type of group did not have a significant effect on in what ways query reformulations would affect the search results.

To sum up, since there is no statistically significant association between groups on the preferred reason for reformulating queries except for one question related to the library search task (i.e., when do you think you needed to reformulate your queries?), it is useful to see overall descriptive statistics of query reformulation experience for both groups at the same time. Detailed descriptive information about each task is provided below.

When do you think you needed to reformulate your queries?

For this question, after completing the virus search task 39.2% answered that they reformulated queries when they wanted to find more relevant information, while 37.3% answered that they reformulated previous queries when their search results were not satisfactory (See Figure 4.12). Based on this information, we can see that relevance and satisfaction are the two most often stated reasons for reformulating queries.

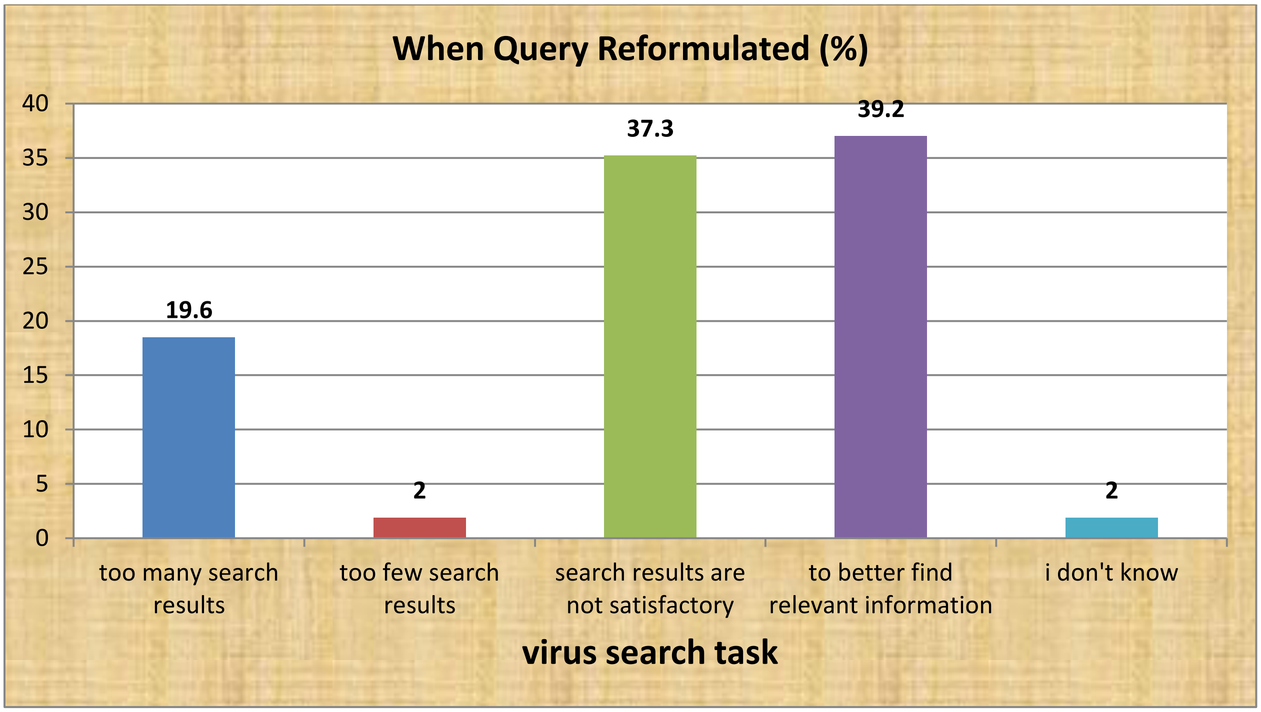

Figure 4.12. Summary of descriptive statistics after the virus search task for the question *when do you think you needed to reformulate your queries?*

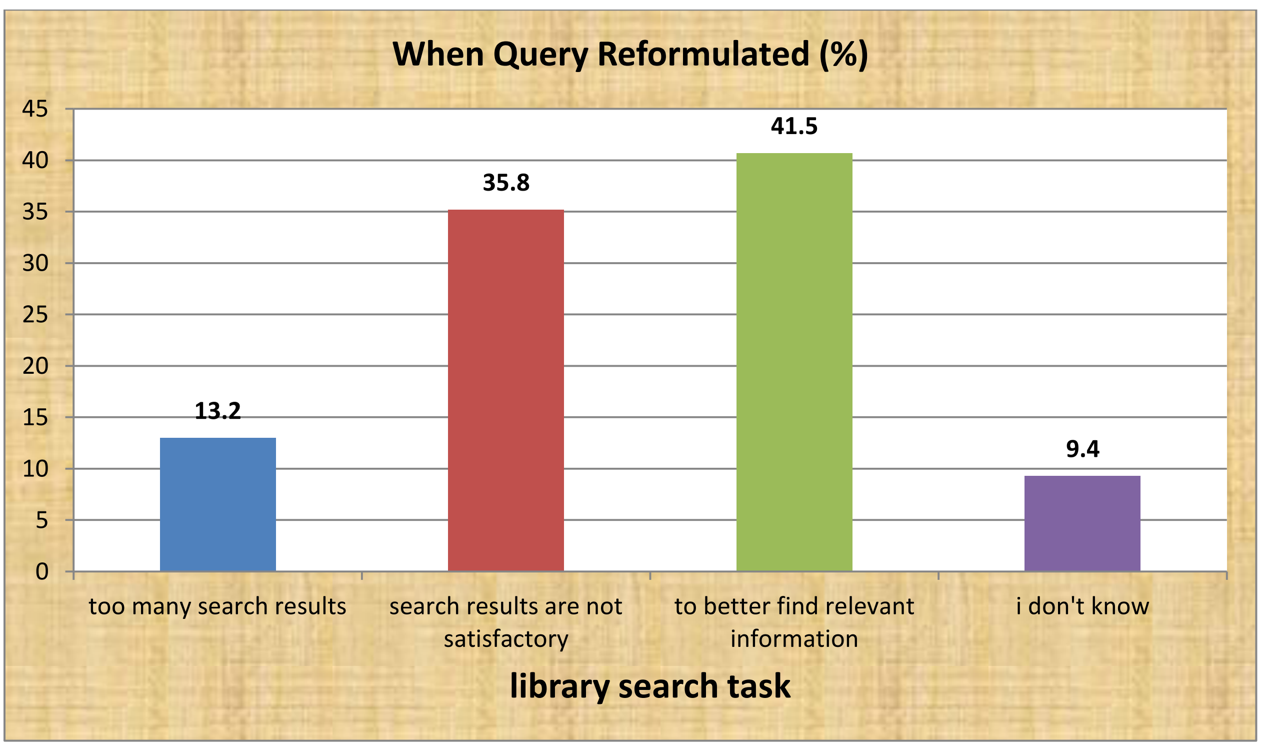

Figure 4.13. Summary of descriptive statistics after the library search task for the question *when do you think you needed to reformulate your queries?*

As shown in Figure 4.13, following completion of the library search task 41.5% answered that they reformulated queries when they wanted to better find relevant information, while 35.8% answered they reformulated previous queries when their search results were not satisfactory. Based on this information, we can see that relevance and satisfaction are the two most often stated reasons for reformulating previous queries. This is consistent with the previous search task.

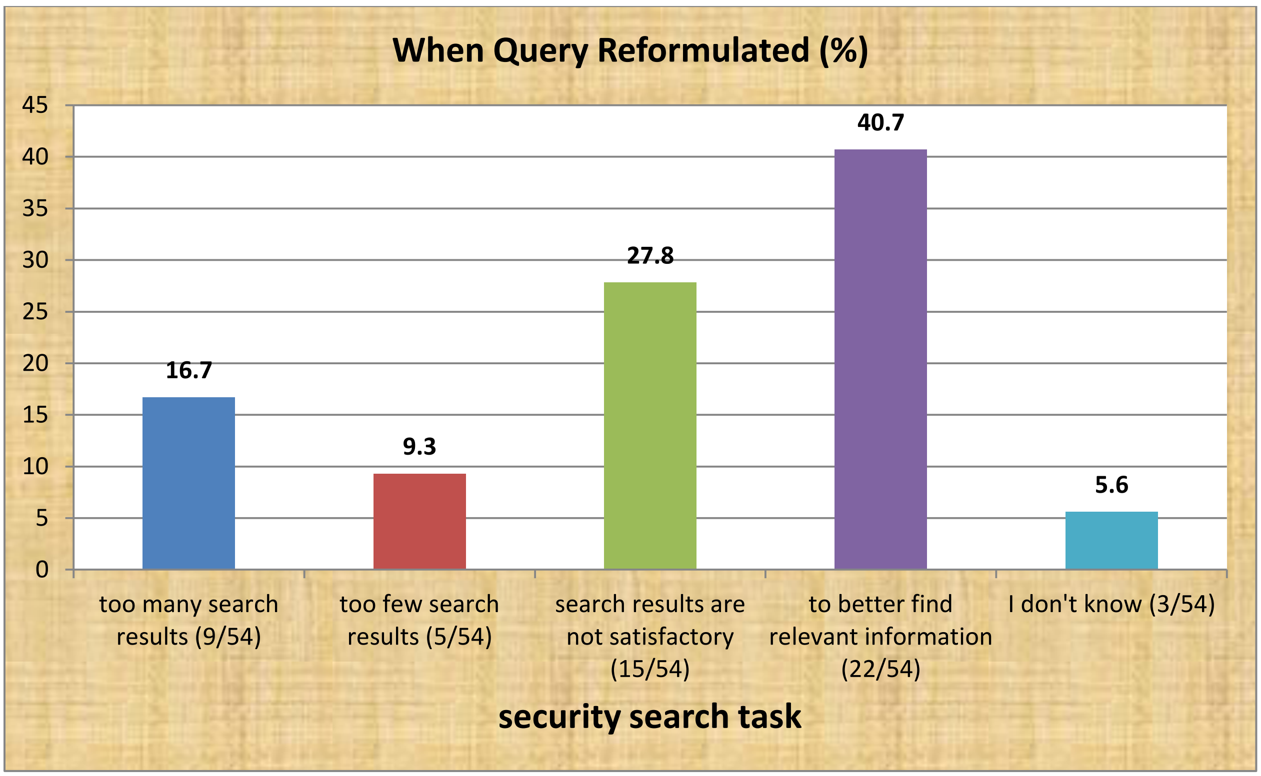

Figure 4.14. Summary of descriptive statistics after the security search task for the question *when do you think you needed to reformulate your queries?*

As shown in Figure 4.14, following the security search task 40.7% responded that relevance was their reason for reformulating queries, while 27.8% responded that performance was their reason for reformulating queries. This is consistent with the two previous tasks.

Why do you think you reformulated your previous queries?

For the second question in section2 of post-task questionnaire (See Figure 4.15), after completing search task1 39.2% answered that they reformulated queries because they wanted to increase precision, while 27.5% answered that they reformulated queries because they wanted to increase relevance. Based on this information, we can see that relevance and precision are two most often stated reasons for reformulating queries.

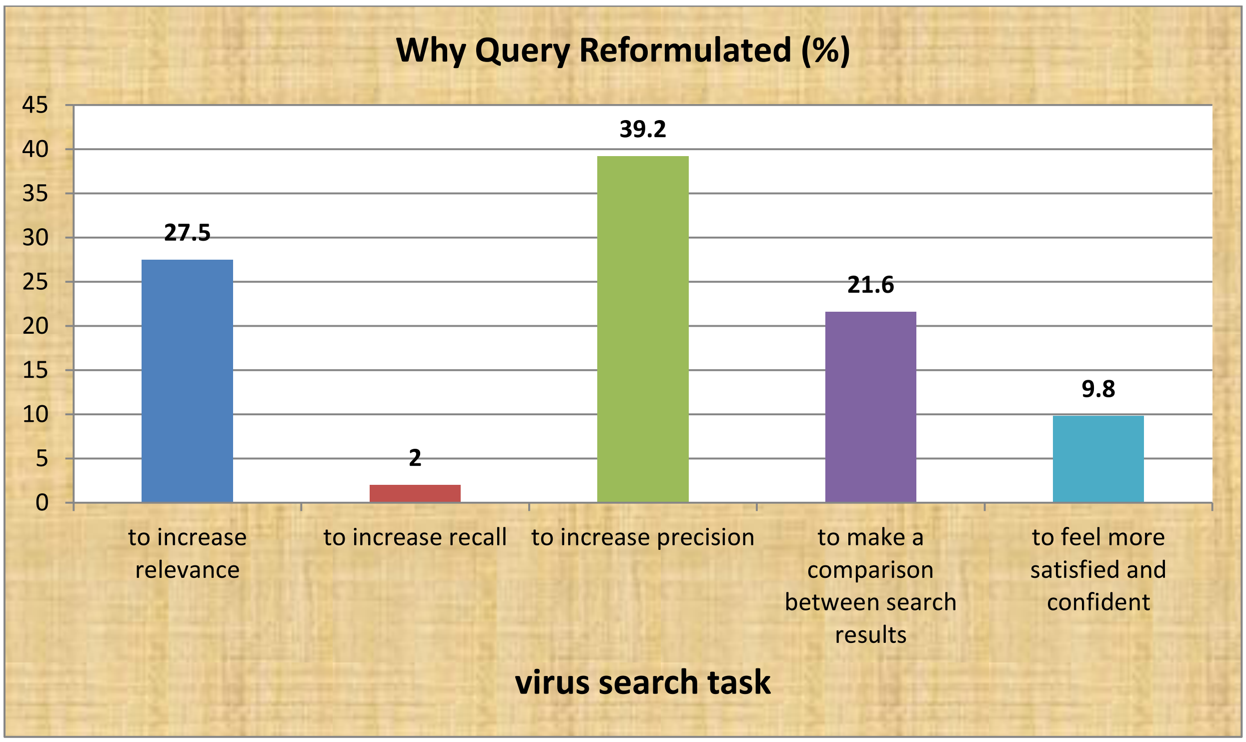

Figure 4.15. Summary of descriptive statistics after the virus search task for the question *why do you think you reformulated your previous queries?*

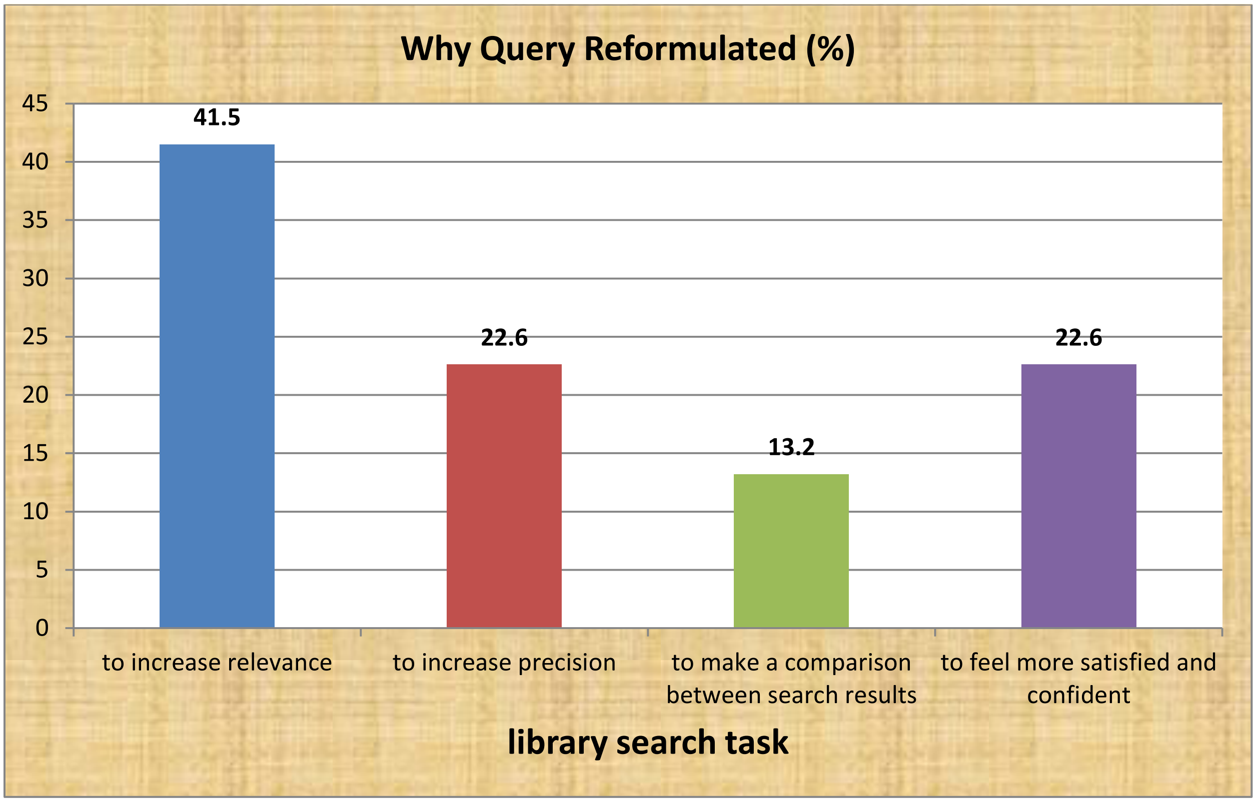

Figure 4.16. Summary of descriptive statistics after library search task for the question *why do you think you reformulated your previous queries?*

As shown in Figure 4.16, after completing search task2 41.5% answered that they reformulated queries because they wanted to increase relevance, while 22.6% answered that they wanted to increase precision and to make a comparison between search results. This information further indicates that relevance and precision are the two most often stated reasons for query reformulation, and that satisfaction and confidence are also important.

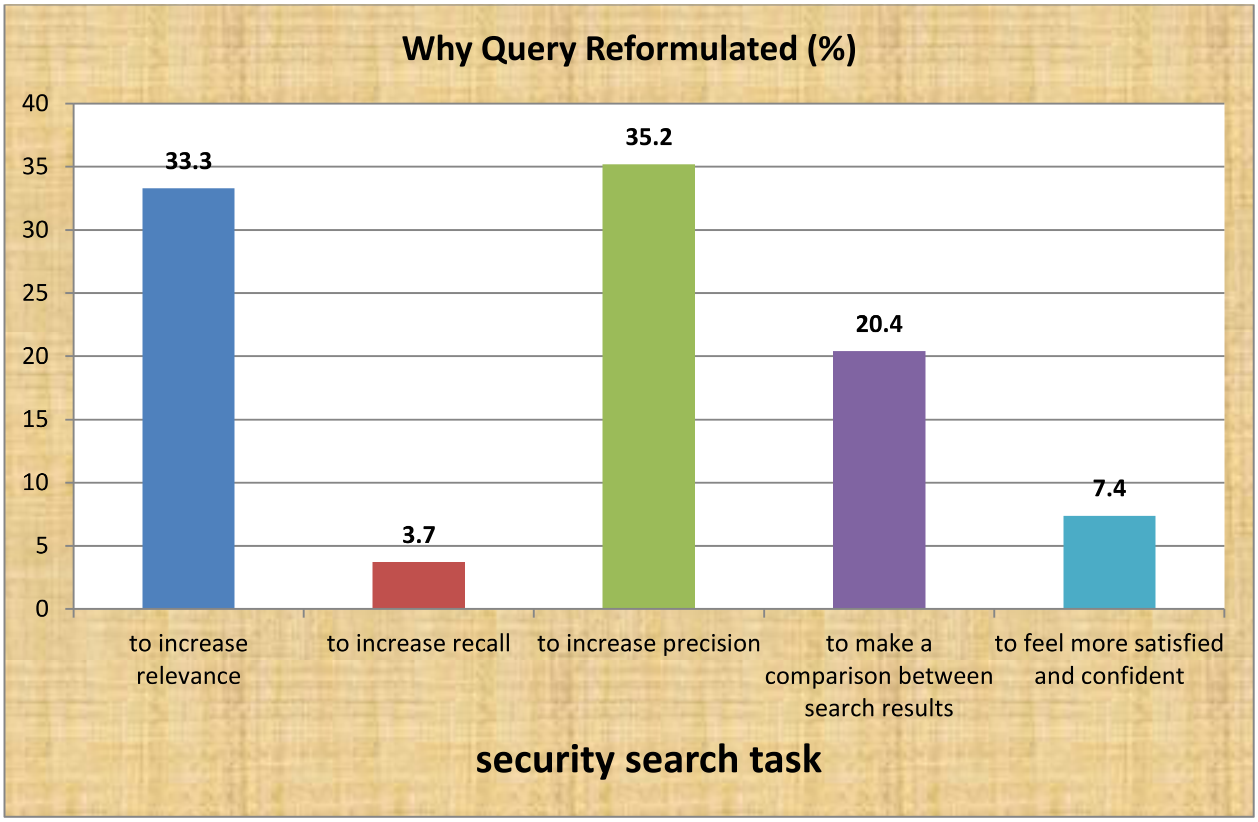

Figure 4.17. Summary of descriptive statistics after the security search task for the question *why do you think you reformulated your previous queries?*

As shown in Figure 4.17, after completing search task3 35.2% answered that they reformulated queries because they wanted to increase precision, while 33.3% answered that wanted to increase relevance. Interestingly, 20.4% responded that they reformulated queries because they wanted to make a comparison between search results. This information indicates that relevance and precision are again the two most often stated reasons for reformulating queries, and comparing search results is also important.

In what ways did the query reformulations affect the search results throughout a search?

For this question, after completing search task1 35.3% answered they thought query reformulations affected the search results in a way that improved the search results, while 31.4% answered in a way that narrowed the search results (See Figure 4.18). It is interesting that 17.6% answered that query reformulation helped inspire new thinking and ideas. This implies that the subjects expected to learn new information during the search process. Based on this information, we can see that improving and narrowing the search results are the two most often stated reasons that query reformulation affects a search.

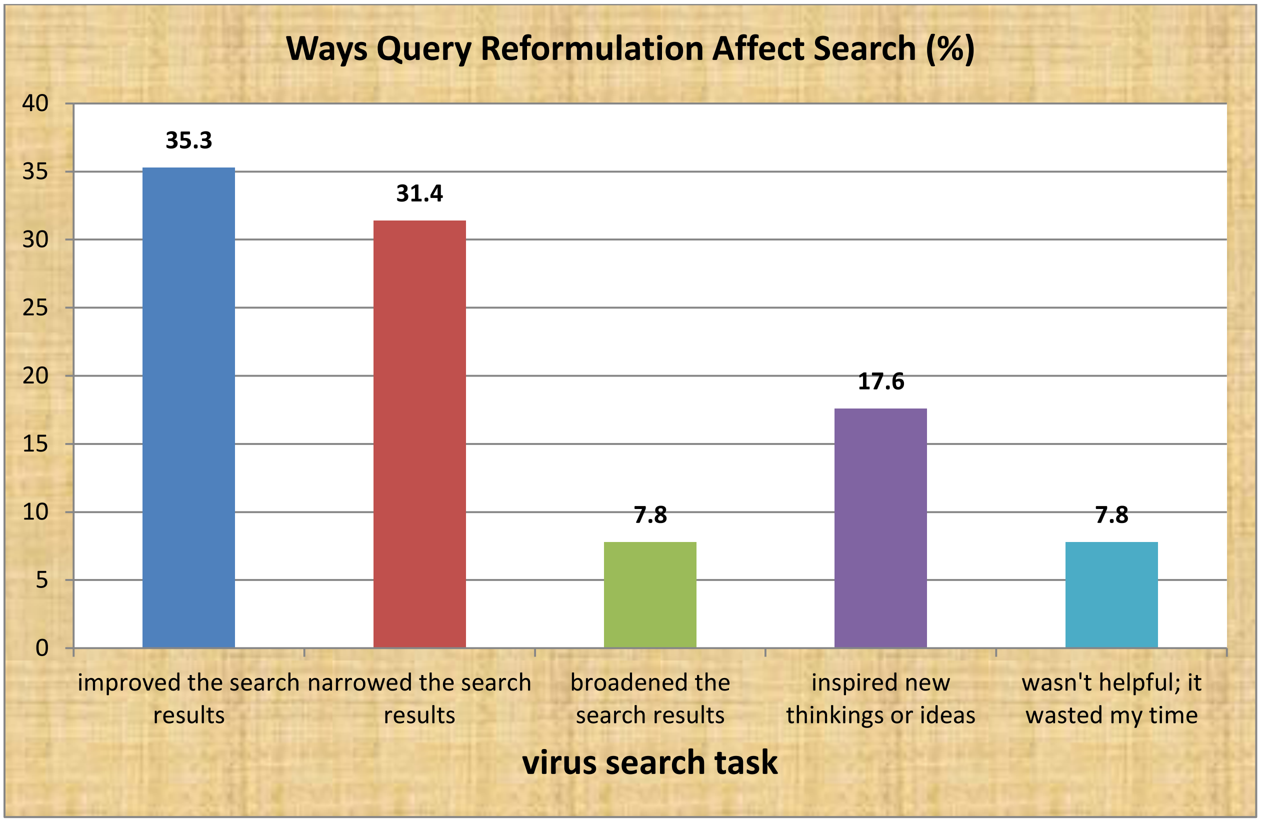

Figure 4.18. Summary of descriptive statistics after virus search task for the question *in what ways did the query reformulations affect the search results throughout a search?*

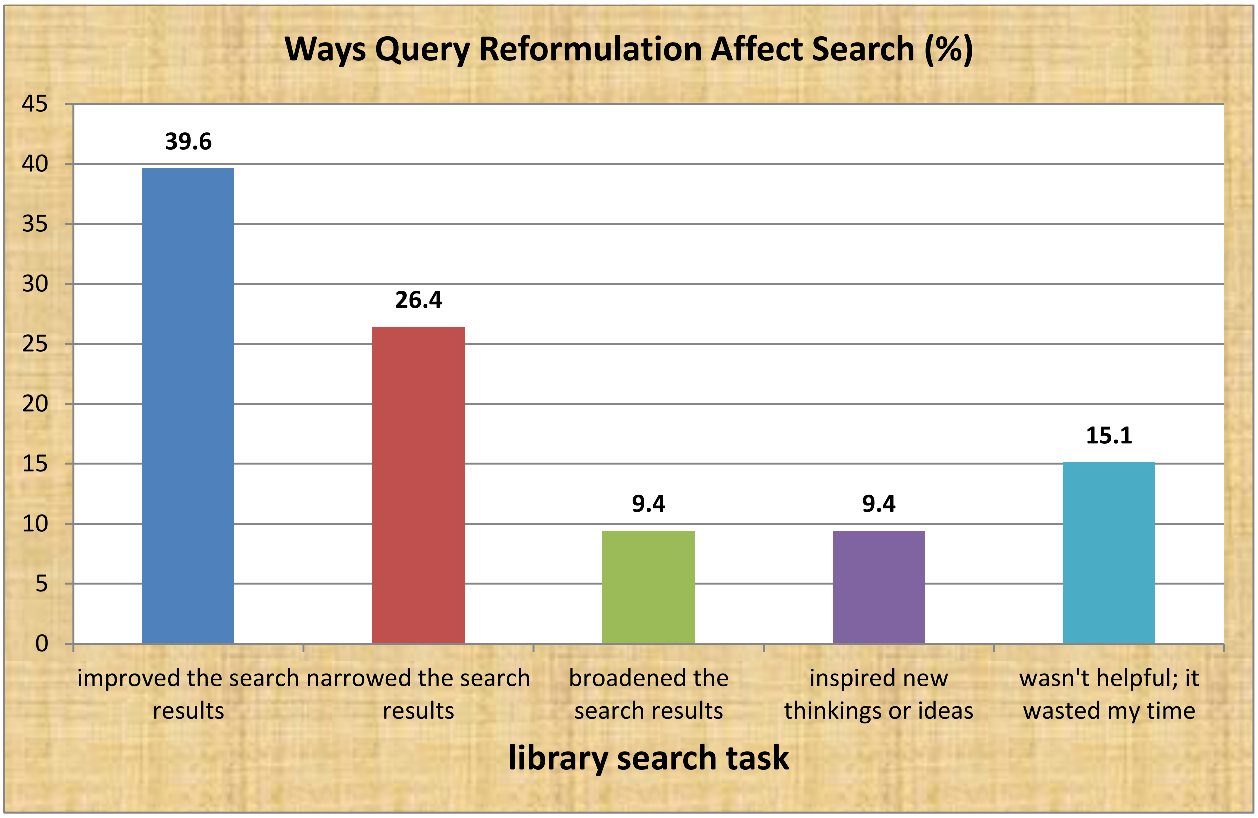

Figure 4.19. Summary of descriptive statistics after library search task *in what ways did the query reformulations affect the search results throughout a search?*

As shown in Figure 4.19, after completing the search task2 39.6% answered that they thought query reformulations affected the search results in a way that improved the search results, while 26.4% answered in a way that narrowed search results. It is interesting to note that 15.1% answered that query reformulation was not helpful and wasted their time. Based on this information, we can see that improving and narrowing the search results are two most often stated reasons for the effect of query reformulations.

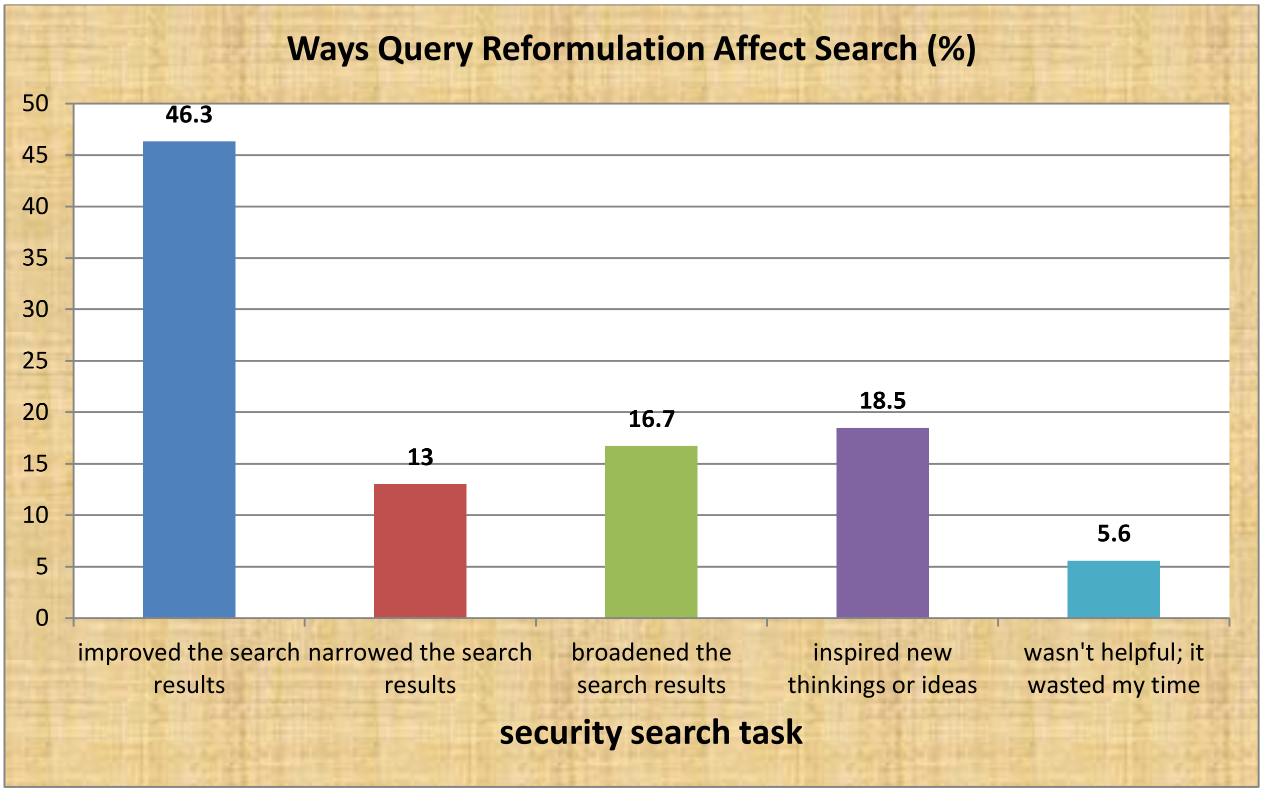

Figure 4.20. Summary of descriptive statistics after security search task for the question *in what ways did the query reformulations affect the search results throughout a search?*

As shown in Figure 4.20, after completing the search task3 46.3% answered they thought query reformulations affected the search results in a way that improved the search results, while 13.0% answered that they thought query reformulation affected search results in a way that narrowed the search results. Interestingly, 16.7% thought that query reformulations affected the search results in a way that broadened the search results, while 18.5% thought that query reformulations inspired new thinking or ideas. This could mean that query reformulations might have not only broadened the search depending upon the quality of query reformulations, but may also have provided opportunities to learn and come up with new ideas. Nevertheless, the most often stated reason for query reformulation was to improve the search results.

Summary

This chapter discussed data analysis and results. In the next chapter, the findings of this study regarding the effect of cognitive load on the propensity for query reformulation behavior will be used to provide insightful and practical implications not only for the design of information retrieval systems, but also for user interface design and website development in e-business. The next chapter concludes the study and addresses its implications. The key research question and three sub-questions were answered in this chapter as follows.

There is a significant difference between the two groups in terms of the propensity for query reformulations. Searchers who experienced cognitive manipulation made fewer queries than searchers who did not experience it. There is also a significant difference between the two groups in terms of the propensity of query reformulations when mental demand, temporal demand, or frustration increases. Searchers who experienced mental demand, temporal demand, or frustration manipulation generated fewer query reformulations than searchers who did not experience it.

There is a significant difference between the two groups in terms of the mental demand, temporal demand, or frustration score in NASA-TLX. Searchers who experienced cognitive load manipulations scored higher on NASA-TLX assessment than searchers who did not experience cognitive load manipulations.

CHAPTER FIVE

CONCLUSION

This study examined the effect of cognitive load on the propensity for query reformulation behavior while using the Google search engine for information retrieval. Three components that contribute to cognitive load were manipulated to demonstrate that the likelihood that a searcher will make query reformulations decreases as the cognitive load increases. This implies that increased cognitive load will interfere with the searcher's propensity to make more queries during the search process. This study did not examine what happens to overall performance and satisfaction with a search when the searcher experiences increased cognitive load, although this would be an interesting area for future study.

To increase the effectiveness and efficiency of the search process, efforts to reduce cognitive load should be a focus of information retrieval system and user interface design. Participants in the experimental group made fewer query reformulations when subjected to mental demand, temporal demand, or frustration manipulations than participants in the control group. This does not mean that more reformulated queries are better than fewer query reformulations; rather in this study query reformulations were considered as one important interaction between the system and the user. Without user queries, a system is useless. Therefore, understanding human cognition in information retrieval through the study of cognitive load assessments can provide useful information for designing and developing effective and efficient information retrieval systems.

As people are engaged in looking for information on the web for their daily routines, it is also important to provide interactions between searchers and systems that are as easy and intuitive as possible. However, lack of understanding caused by invisible impediments or sources that contribute to cognitive load might present substantial problems for both system developers and searchers. Therefore, understanding the effect of cognitive load on the propensity for query reformulation behavior can help develop and refine theories of human information seeking behavior on the web and in general.

Implications for designing information systems, user interfaces, or websites, and practice

Given the current nature of the web, higher precision and lower recall is desirable because one can become overwhelmed with millions of hits to assess for relevance. As the Web continues to evolve, the structure of the Web contents becomes more complex. However, the lack of understanding the major impediments to the user within that structure may become a substantial problem to both system developers and users and it is certainly important to understand how the cognitive load of searchers works during the information search process in order to improve the interaction between the systems and the users. Therefore information system designers and developers should explore the user's cognitive load level to determining what factors to focus on to improve the interaction so that they can then take advantage of the results of cognitive load measurements and work to redesign the problem areas. A recent study (Yurko, Scerbo, Prabhu, Acker, & Stefanidis, 2010) found that increased mental workload results in inferior task performance and a higher likelihood of errors. They claimed that individuals with high workloads are more prone to error and their NASA-TLX scores accurately reflected workload differences during simulator training.

By exploring aspects of cognitive activity when searchers engage in finding relevant information on the web, researchers can help design and develop better and more usable information search systems, user interfaces, and websites. This study examined query reformulation behavior to explore how imposed cognitive load might impede the search experience during information search process. Query reformulation behavior is an important interaction between a system and a searcher; this study lays the groundwork for further research and sheds light on previously unexplored aspects of query reformulation behavior related to the effect of cognitive load. Moreover, this study provides significant insights to system designers and information professionals, including the following:

1) Searchers are NOT likely to make more queries when they are experiencing higher/increased cognitive load than lower/reduced cognitive load;
2) System developers should pay attention to variations in the cognitive load of searchers during the search process;

3) System developers should improve query formulation support and education so that searchers can utilize them to achieve search satisfaction;
4) Information professionals should pay attention to the cognitive load of searchers so that they can achieve better search satisfaction and performance;
5) Information professionals should pay attention to query reformulation behavior.

Reducing cognitive load

Some guidelines and recommendations for designing user interfaces have been elaborated to help designers and developers take into account the limitation of working memory capacity. Shneiderman's Eight Golden Rules of Interface Design (Schneiderman & Plaisant, 1998; Shneiderman & Plaisant, 2005) include the reduction of short-term memory load. Other rules include striving for consistency; enabling frequent users to use shortcuts; offering informative feedback; designing dialogs to yield closure; offering error prevention and simple error handling; permitting easy reversal of actions; and supporting an internal locus of control (Schneiderman & Plaisant, 1998; Shneiderman & Plaisant, 2005).

Mandel also discusses principles that could be implemented when designing and developing information systems to reduce users' memory and cognitive load (Mandel, 2002):

- Relieve short-term memory;
- Rely on recognition, not recall;
- Provide visual cues;
- Provide interface shortcuts;
- Promote an object-action syntax;
- Use real-world metaphors;
- Use progressive disclosure; and,
- Promote visual clarity.

However, research concerning the effect of cognitive load on the propensity for query reformulation behaviors in terms of information retrieval is sparse. The results of this study support recommendations for reducing cognitive load to improve information search experience in information systems through better support and education. For example, if Google were to provide auto-generated search queries or a preview feature, searchers could get a better idea of what they should do next. Automated aids such as "did you mean~?" or "question answering," or "auto-generated keyword support," which give the searcher the option to choose suggested keywords from a system may help reduce mental demand in the search process. Some of these recommendations may not be directly relevant to the design of information retrieval systems. Still, it is important to better understand communication between users and systems.

Schmutz, et al. (2009) studied the effect of cognitive load in eCommerce web applications for four online book stores: amazon.ch, buch.ch, book.ch, and buchhaus.ch. The purpose of their research was to examine how cognitive load might influence preference, performance, and satisfaction. They found that cognitive load significantly differed among the shops and found strong correlations between NASA-TLX score, primary task completion time, preference, and satisfaction. Queries can play a large role in initiating and continuing interactions between a system and user. This could be important for e-business in that one more query could lead to another item to sell. Amazon.com provides features to reduce queries such as "one-click" providing hyperlinks for "*Customers Who Bought This Item Also Bought*," "*Frequently Bought Together,*" "*Customer Review*," and "*Look Inside*." This might imply that designers have already considered reducing cognitive load in that queries could be replaced by clicking pre-established options.

One major concern regarding cognitive load in information searching is frustration. For example, smartphone users may have a difficult time in making queries because smartphone input devices are smaller than a computer keyboard. The keys are too small and too closely spaced for computer keyboard and touch-type trained fingers. To improve this kind of interaction between a user and a system, speech recognition and touch screen have been researched and implemented. While those features are currently implemented and used by searchers, speech recognition search systems could increase searching and reformulations if the system is designed well. Small keys are physically and cognitively demanding. If a system's physical demands are high, then a user might feel exhausted and give up on a task. Frustration might lead to low satisfaction.

Hearst (2009) states that "in order to show users helpful alternatives, researchers have developed several techniques to try to aid in the query reformulation process, although existing tools may not be sophisticated enough to aid the user with the information need" (p.142). Hearst

(2009) also discusses some interface technologies to support query reformulation and the ways in which users interact with them.

1) Spelling suggestions and corrections: Search transaction logs show that searchers often make mistakes on queries like spelling or typographical errors (Cucerzan & Brill, 2004). Hearst (2009) mentions that "one important query reformulation tool is spelling suggestions or corrections" (p.143). The Google search engine corrects misspelled queries and provides the alternative "search instead for " as shown in Figure 5.1.

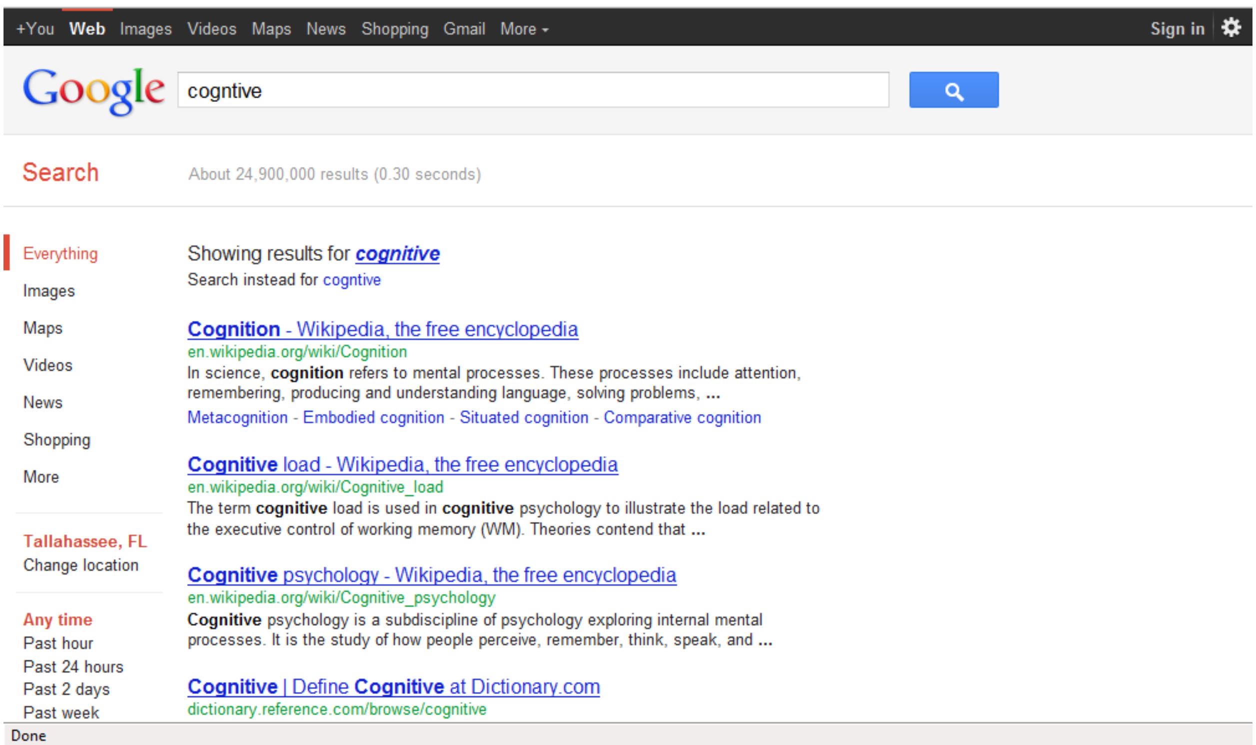

Figure 5.1. Showing results for misspelled query from Google ("Search instead for")

2) Automatically suggested term: Hearst (2009) also mentions that some "query reformulation aids are automatically suggested term refinements and expansions. Spelling correction suggestions are also query reformulation aids, but the phrase term expansion is usually applied to tools that suggest alternative words and phrases" (p.143). White et al (2007) showed that query refinement suggestions in search engines are generally helpful. Google automatically generates suggested terms when a searcher begins inputting a query

as shown in Figure 5.2. Bing offers related searches as shown in Figure 5.3.

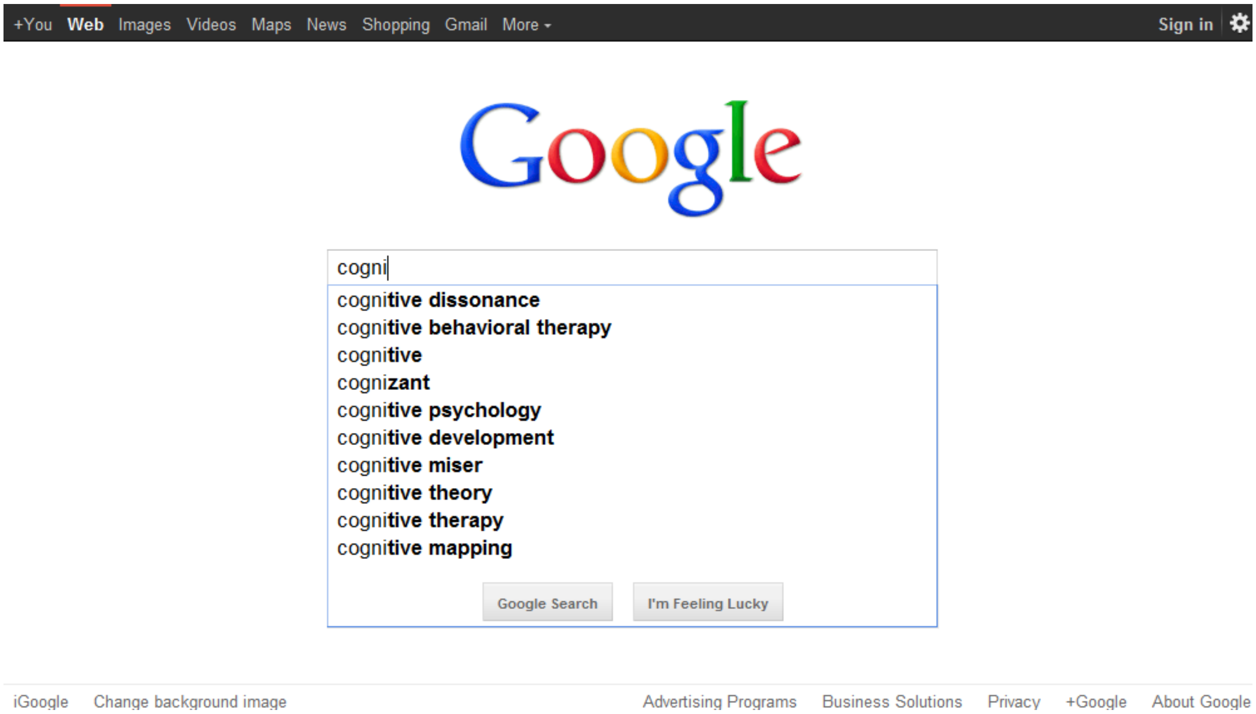

Figure 5.2 Automatic suggested terms from Google

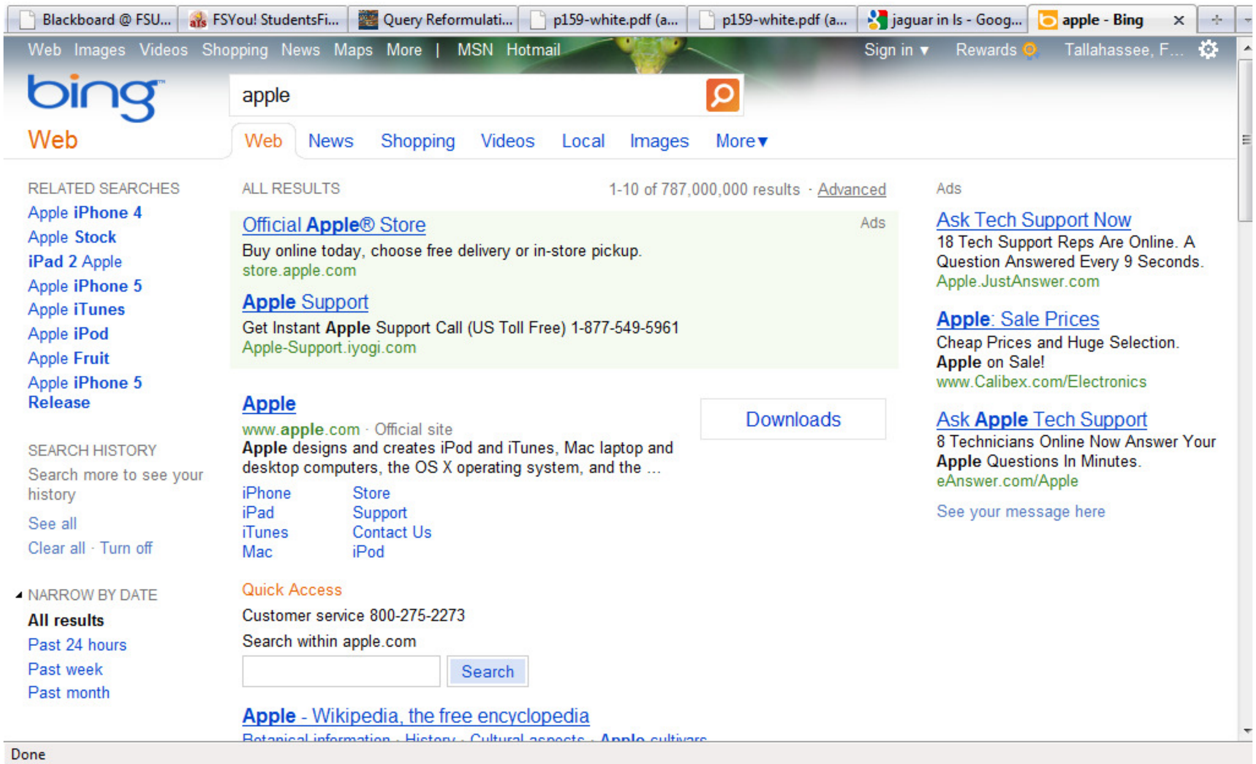

Figure 5.3. Related searches from the Bing search engine

In conclusion, the evaluation of information retrieval systems has focused on performance in finding relevant information to meet searchers' information needs. Furthermore, it is equally important and expected that the evaluation of human cognitive load in the search process should be considered to better facilitate human search experience when designing information retrieval systems and user interfaces. As discussed earlier, the results of this study show that higher cognitive load imposed on a searcher can interfere with end users' search experience in query reformulation. Overall, the implications of this study concerning the effect of cognitive load on the propensity for query reformulation as one type of interaction between a user and a system can be used to provide insights into developing information retrieval systems and enhance user interfaces.

Implications for future research

Although cognitive load theory has been rigorously studied in psychological, educational and instructional settings, cognitive load is just now gradually being considered in the design and development of information systems and user interfaces. Cognitive load has been studied rigorously in disciplines of Education, Cognitive Psychology, and Industrial Engineering, but Library and Information Studies lacks research on this matter. Previous research on information retrieval focused on the performance of information search systems, and studies of query formulation behavior have examined patterns and transition of queries. It will be interesting to see if cognitive load correlates with performance, satisfaction, preference, time, and other components that might be important in LIS. Other sources of cognitive load than mental demand, temporal demand, and frustration should also be carefully considered.

This research provides a platform for future studies to examine the dependent variable of time or duration estimation as a cognitive load measurement. Block, Hancock, and Zakay (2010) claim that time estimation has been shown in several experiments to be a reliable and valid measure of cognitive load. This study also paves the way for future research that can employ different data collection techniques such as eye-tracking and physiological assessment of cognitive load in information seeking. Future studies might require only one search task instead of three that might induce cognitive fatigue due to carryover effect.

Although this study did not test the usability of Google's search engine, the methodology used for this study can be utilized for usability tests because analyzing the components of cognitive load in the NASA-TLX workload assessment can provide correlative evidence of usability. In this study, dual-task and NASA-TLX served well in terms of examining cognitive load. Other multi- or single methods could be adapted for the future methodologies using the NASA-TLX workload assessment to better understand cognitive processes and cognitive load when searching for information. In terms of target populations, the study population may be expanded to different discipline majors to examine whether similarities or differences exist.

Summary

This study shows that a decrease in the number of query reformulations is associated with increased cognitive load. Increased cognitive load scores are associated with increased workload manipulations of mental demand, temporal demand, and frustration. Search query reformulation behavior is affected by cognitive load and these effects differ depending upon which manipulations are introduced.

1) When subjects received increased mental demand while performing a search task, they formulated fewer queries than those who did not experience increased mental demand.
2) When subjects had limited time allocated while performing a search task, they also formulated fewer queries than those who did not have time limitations.
3) When subjects were forced to use an unfamiliar on-screen keyboard to increase frustration while performing a search task, they also formulated fewer queries than those who were not forced to use it.

The NASA-TLX workload assessment seems to be a reliable and valid way to assess cognitive load in this study.

The results of this study stress the importance of the effect of cognitive load in the propensity for query reformulation behavior. This study indicated that increased cognitive load in information retrieval on the Web can result in significant difference in the propensity for query reformulation. This study also demonstrates increased cognitive load imposed on searchers considerably increases their NASA-TLX workload assessment scores.

APPENDIX A

Pre-Task Questionnaire

The purpose of this questionnaire is to collect demographic information, information about your computer skills, Web searching skills, and prior experience with the Google search engine.

The questionnaire is composed of the following sections:

I. Section 1: Demographics
II. Section 2: Computer knowledge, Web search experience, and prior experience with the Google Web search engine

Your answers are extremely important to the reliability and validity of this research, so please answer all the questions.

Pilot tests indicate that it will take less than five minutes to complete this survey. Responses will be reported only in aggregates and your identity will not be disclosed.

If you have any questions as you complete the questionnaire, the researcher will be available to answer them.

Thank you very much for your participation.

Section 1: Demographics

Please choose the answer that best describes you.

1. Gender
   A. Male
   B. Female

2. Date of Birth (e.g., 8/1/1981) ________________________

3. Which degree are you currently pursuing? (Choose one that best describes your situation.)
   A. BA or BS
   B. MS or MA
   C. Ph.D
   D. Specialist (i.e., post-master's)
   E. Others (e.g.,certificate or other non-degree program)

Section 2: Computer Knowledge, Web search experience, and Prior Experience with the Google search engine

Please answer the following questions to explain your computer knowledge, Web search experience and skills, and prior experience with the Google search engine.

1. How would you rate your knowledge and experience with computers?
   (Circle one that best describes you.)

Very low	low	average	high	Very high
1	2	3	4	5

2. How would you rate your knowledge and experience with Web searching?
   (Circle one that best describes you.)

Very low	low	average	high	Very high
1	2	3	4	5

3. How would you rate your search skills on the Web in general?
   (Circle one that best describes you.)

Very low	low	average	high	Very high
1	2	3	4	5

4. How frequently do you use the *Google* search engine (http://www.google.com)?
   (Circle the answer that best describes the frequency of your use)

Seldom	Weekly	At least once a week	Daily	More than once a day
1	2	3	4	5

5. In previous searches using the Google search engine, how many times did you attempt to refine your search?
   (Circle the response that best describes you.)

Never	Rarely	About 1 time	About 2 times	More than 3 times
1	2	3	4	5

6. In the past, how successful would you say your Google searches have been?
   (Circle one that best describes your past search success.)

Never	Rarely	Sometimes	Usually	Always
1	2	3	4	5

Next, I will introduce you to the experiment protocol and you will participate in a practice session prior to beginning the experiment.

APPENDIX B

Post-Task Questionnaire

Section 1: Cognitive Load

1. **Mental demand**

   Rate your level of mental and perceptual activity (i.e., thinking, looking, searching, or remembering) while performing the search task. How mentally demanding was it?

   (Circle the numeral that best represents your rating)

Extremely low	Very low	In between very low and medium	Medium	In between medium to very high	Very high	Extremely high
1	2	3	4	5	6	7

2. **Physical demand**

   Rate your level of physical activity (i.e., clicking, scrolling, or typing) while performing the search task. How physically demanding was it?

   (Circle the numeral that best represents your rating)

Extremely low	Very low	In between very low and medium	Medium	In between medium to very high	Very high	Extremely high
1	2	3	4	5	6	7

3. **Temporal demand**

   Rate the level of perceived time pressure to complete the search task? How hurried or rushed was the pace of the task? How much pressure did you feel to complete the task quickly?

   (Circle the numeral that best represents your rating)

Extremely low	Very low	In between very low and medium	Medium	In between medium to very	Very high	Extremely high

				high		
1	2	3	4	5	6	7

4. **Performance**

Rate your level of performance in performing the search task. How successful were you in accomplishing what you were asked to do? How satisfied were you with your performance? (Circle the numeral that best represents your rating)

Extremely high (perfect)	Very high	In between very high and medium	Medium	In between medium to very low	Very low	Extremely low (failure)
1	2	3	4	5	6	7

5. **Effort**

Rate your level of effort as you performed the task. How hard did you have to work (mentally and physically) to accomplish it?

(Circle the numeral that best represents your rating)

Extremely low	Very low	In between very low and medium	Medium	In between medium to very high	Very high	Extremely high
1	2	3	4	5	6	7

6. **Frustration**

Rate your level of frustration as you performed the task. How insecure, discouraged, irritated, stressed, and annoyed did you feel?

(Circle one numerical number)

Extremely low	Very low	In between very low and medium	Medium	In between medium to very high	Very high	Extremely high
1	2	3	4	5	6	7

Section 2: Overall Query Reformulation Experience

1. When do you think you needed to reformulate your queries? **(Circle one that best describes you.)**
    A. When there were too many search results
    B. When there were too few search results
    C. When the search results are not satisfactory
    D. To better find relevant information
    E. I don't know

2. Why do you think you reformulated your previous queries? **(Circle one that best describes you.)**
    A. To enhance relevance
    B. To enhance recall
    C. To enhance precision
    D. To make a comparison between the search results of queries
    E. To feel more satisfied and confident

3. In what ways did the query reformulations affect the search results throughout a search? **(Circle one that best describes you.)**
    A. It improved the search results
    B. It narrowed the search results
    C. It broadened the search results
    D. It inspired new thinking/ ideas
    E. It wasn't helpful; it wasted my time

4. Do you have any other comments about this project?

Thank you very much!

APPENDIX C

Introduction (Control Group)

This research project is designed to investigate the effect of cognitive load on the propensity for query reformulation behavior using Google.

You will be exposed to three search tasks in order. At the completion of each task, you should complete a short survey questionnaire.

I would like to encourage you to *complete your search tasks in order to find the information that meets the goals of the task*. If you are satisfied with your search results, you can stop the task. However, if for some reason you are not comfortable with the task given during your search process, you can quit your search at any time.

During the tasks, you may not ask questions about how to conduct the search, but you may ask questions about technical problems such as the network connection, system malfunction, and the like.

Thank you very much for your participation!

Control group

Introduction (Experimental Group)

This research project is designed to investigate the effect of cognitive load on the propensity for query reformulation behavior using Google.

You will be exposed to three search tasks in order. At the completion of each task, you should complete a short survey questionnaire.

I would like to encourage you to *complete your search tasks in order to find the information that meets the goals of the task*. If you are satisfied with your search results, you can stop the task. However, if for some reason you are not comfortable with the task given during your search process, you can quit your search at any time.

During the tasks, you may not ask questions about how to conduct the search, but you may ask questions about technical problems such as the network connection, system malfunction, and the like.

Thank you very much for your participation!

Experimental group

Practice Task

Practice task

- Suppose you are a researcher exploring why people use social networking tools (i.e., twitter, facebook, youtube, blog, etc.)? What are the affective factors that cause people to join and use social networking tools listed above? What are advantages and disadvantages of using these social networking tools?

Please write your best answer or URL below.

EG1A (Experimental Group 1 Task order A)

Task1

(Virus-Mental): You will be given 8 digits number such as “32841792” before you start the search task. It is very important for you to memorize and recall this number after you complete the search task. You will be presented this number before you start the search task when you are ready. The experimenter will ask you this number when you stop the search task completely.
Task: Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can’t find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please write down 8 digits number you were given:

(____________________________________)

Please complete a post-survey questionnaire and turn to the next page!

Task2

(Library-Temporal): You are only given 5 minutes to complete this search task. Please complete the search task within 5 minutes. You can consult the egg timer to manage your time to complete the task. The experimenter will set up the egg timer before you start the task.

Task: Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies will you bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task3

(Security-Frustration): Given the apparatus, On-Screen Keyboard, please complete the search task. Please know that you are only allowed to use On-Screen Keyboard and mouse while completing this search task.

Task: Suppose you are a web developer who is in charge of client's website security to prevent hacking and malicious attacks to the system. What kinds of prevention methods you will propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire and Thank you very much!

EG1B (Experimental Group 1 Task order B)

Task1

(Virus-Temporal): You are only given 5 minutes to complete this search task. Please complete the search task within 5 minutes. You can consult the egg timer to manage your time to complete the task. The experimenter will set up the egg timer before you start the task.
Task: Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task2

(Library - Frustration): Given the apparatus, On-Screen Keyboard, please complete the search task. Please know that you are only allowed to use On-Screen Keyboard and mouse while completing this search task.

Task: Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task3

(Security - Mental): You will be given 8 digits number such as "32841792" before you start the search task. It is very important for you to memorize and recall this number after you complete the search task. You will be presented this number before you start the search task when you are ready. The experimenter will ask you this number when you stop the search task completely. Task: Suppose you are a web developer who is in charge of client's website security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please write down 8 digits number you were given:

(______________________________________)

Please complete a post-survey questionnaire!

EG1C (Experimental Group 1 Task order C)

Task1

(Virus - Frustration): Given the apparatus, On-Screen Keyboard, please complete the search task. Please know that you are only allowed to use On-Screen Keyboard and mouse while completing this search task.

Task: Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task2

(Library - Mental): You will be given 8 digits number such as “32841792” before you start the search task. It is very important for you to memorize and recall this number after you complete the search task. You will be presented this number before you start the search task when you are ready. The experimenter will ask you this number when you stop the search task completely.

Task: Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please write down 8 digits number you were given:

(_____________________________________)

Please complete a post-survey questionnaire!

Task3

(Security - Temporal): You are only given 5 minutes to complete this search task. Please complete the search task within 5 minutes. You can consult the egg timer to manage your time to complete the task. The experimenter will set up the egg timer before you start the task.

Task: Suppose you are a web developer who is in charge of client's website security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

EG2A (Experimental Group 2, Task order A)

Task1

(Library - Mental): You will be given 8 digits number such as “32841792” before you start the search task. It is very important for you to memorize and recall this number after you complete the search task. You will be presented this number before you start the search task when you are ready. The experimenter will ask you this number when you stop the search task completely.

Task: Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please write down 8 digits number you were given:

(______________________________________)

Please complete a post-survey questionnaire!

Task2

(Security - Temporal): You are only given 5 minutes to complete this search task. Please complete the search task within 5 minutes. You can consult the egg timer to manage your time to complete the task. The experimenter will set up the egg timer before you start the task.
Task: Suppose you are a web developer who is in charge of client's website security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task3

(Virus - Frustration): Given the apparatus, On-Screen Keyboard, please complete the search task. Please know that you are only allowed to use On-Screen Keyboard and mouse while completing this search task.

Task: Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

EG2B (Experimental Group 2, Task order B)

Task1

(Library-Temporal): You are only given 5 minutes to complete this search task. Please complete the search task within 5 minutes. You can consult the egg timer to manage your time to complete the task. The experimenter will set up the egg timer before you start the task.

Task: Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task2

(Security-Frustration): Given the apparatus, On-Screen Keyboard, please complete the search task. Please know that you are only allowed to use On-Screen Keyboard and mouse while completing this search task.

Task: Suppose you are a web developer who is in charge of client's website security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire.

Task3

(Virus-Mental): You will be given 8 digits number such as “32841792” before you start the search task. It is very important for you to memorize and recall this number after you complete the search task. You will be presented this number before you start the search task when you are ready. The experimenter will ask you this number when you stop the search task completely. Task: Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can’t find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please write down 8 digits number you were given:

(_____________________________________)

Please complete a post-survey questionnaire.

EG2C (Experimental Group 1, Task order C)

Task1

(Library - Frustration): Given the apparatus, On-Screen Keyboard, please complete the search task. Please know that you are only allowed to use On-Screen Keyboard and mouse while completing this search task.

Task: Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task2

(Security - Mental): You will be given 8 digits number such as "32841792" before you start the search task. It is very important for you to memorize and recall this number after you complete the search task. You will be presented this number before you start the search task when you are ready. The experimenter will ask you this number when you stop the search task completely. Task: Suppose you are a web developer who is in charge of client's website security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task3

(Virus-Temporal): You are only given 5 minutes to complete this search task. Please complete the search task within 5 minutes. You can consult the egg timer to manage your time to complete the task. The experimenter will set up the egg timer before you start the task.

Task: Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can’t find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

EG3A (Experimental Group 3, Task order A)

Task1

(Security - Mental): You will be given 8 digits number such as "32841792" before you start the search task. It is very important for you to memorize and recall this number after you complete the search task. You will be presented this number before you start the search task when you are ready. The experimenter will ask you this number when you stop the search task completely.

Task: Suppose you are a web developer who is in charge of client's website security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please write down 8 digits number you were given:

(____________________________________)

Please complete a post-survey questionnaire!

Task2

(Virus-Temporal): You are only given 5 minutes to complete this search task. Please complete the search task within 5 minutes. You can consult the egg timer to manage your time to complete the task. The experimenter will set up the egg timer before you start the task.

Task: Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task3

(Library - Frustration): Given the apparatus, On-Screen Keyboard, please complete the search task. Please know that you are only allowed to use On-Screen Keyboard and mouse while completing this search task.

Task: Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

EG3B (Experimental Group 3, Task order B)

Task1

(Security - Temporal): You are only given 5 minutes to complete this search task. Please complete the search task within 5 minutes. You can consult the egg timer to manage your time to complete the task. The experimenter will set up the egg timer before you start the task.

Task: Suppose you are a web developer who is in charge of client's website security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task2

(Virus - Frustration): Given the apparatus, On-Screen Keyboard, please complete the search task. Please know that you are only allowed to use On-Screen Keyboard and mouse while completing this search task.

Task: Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task3

(Library - Mental): You will be given 8 digits number such as "32841792" before you start the search task. It is very important for you to memorize and recall this number after you complete the search task. You will be presented this number before you start the search task when you are ready. The experimenter will ask you this number when you stop the search task completely.

Task: Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please write down 8 digits number you were given:

(______________________________________)

Please complete a post-survey questionnaire!

EG3C (Experimental Group 3, Task order C)

Task1

(Security-Frustration): Given the apparatus, On-Screen Keyboard, please complete the search task. Please know that you are only allowed to use On-Screen Keyboard and mouse while completing this search task.

Task: Suppose you are a web developer who is in charge of client's website security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task2

(Virus-Mental): You will be given 8 digits number such as "32841792" before you start the search task. It is very important for you to memorize and recall this number after you complete the search task. You will be presented this number before you start the search task when you are ready. The experimenter will ask you this number when you stop the search task completely. Task: Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please write down 8 digits number you were given:

(______________________________________)

Please complete a post-survey questionnaire!

Task3

(Library-Temporal): You are only given 5 minutes to complete this search task. Please complete the search task within 5 minutes. You can consult the egg timer to manage your time to complete the task. The experimenter will set up the egg timer before you start the task.

Task: Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

CG1A (Control Group 1, Task order A)

Task1

(Virus): Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task2

(Library): Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task3

(Security): Suppose you are a web developer who is in charge of security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

CG2B (Control Group 2, Task order B)

Task1

(Library): Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task2

(Security): Suppose you are a web developer who is in charge of security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task3

(Virus): Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

CG3C (Control Group 3, Task order C)

Task1

(Security): Suppose you are a web developer who is in charge of security to prevent hacking and malicious attacks to the system. What kinds of prevention methods will you propose to your client to best protect his or her website?

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task2

(Virus): Suppose your computer was attacked and infected by a virus that prevented you from executing any computer program. Furthermore, the anti-virus software you have installed will not work and you cannot install a new anti-virus program; so you can't find out which virus has attacked your computer. Here is more detail:

a. You cannot execute any of the programs installed on your computer.
b. You cannot install any anti-virus software to find and delete the virus that has infected your computer.
c. You cannot login to any application that requires a login process because your login name and password are not working.

Given these facts, find information on the kind of virus that has infected your computer and possible solutions to resolve this issue so that your computer will work properly.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

Task3

(Library): Suppose a local public library near your residence was burned to the ground. Suppose you are the person who is in charge of rebuilding the library computer system, not the library building. What are the approaches and best strategies to perform in this situation? Find the best strategies you will bring up to rebuild the library computer system.

Please write your best answer or URL below.

Please complete a post-survey questionnaire!

APPENDIX D

Human Subject Approval Memorandum

Office of the Vice President For Research

Human Subjects Committee

Tallahassee, Florida 32306-2742

(850) 644-8673, FAX (850) 644-4392

APPROVAL MEMORANDUM

Date: 1/11/2011

To: Kyoungsik Na [*****@my.fsu.edu]

Address: 2100

Dept.: COLLEGE OF INFORMATION

From: Thomas L. Jacobson, Chair

Re: Use of Human Subjects in Research

EXPLORING THE EFFECT OF COGNITIVE LOAD ON THE PROPENSITY FOR QUERY REFORMULATION BEHAVIOR

The application that you submitted to this office in regard to the use of human subjects in the proposal referenced above have been reviewed by the Secretary, the Chair, and two members of the Human Subjects Committee. Your project is determined to be Expedited per 45 CFR Â§ 46.110(7) and has been approved by an expedited review process.

The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals, which may be required.

If you submitted a proposed consent form with your application, the approved stamped consent form is attached to this approval notice. Only the stamped version of the consent form may be used in recruiting research subjects.

If the project has not been completed by 1/6/2012 you must request a renewal of approval for continuation of the project. As a courtesy, a renewal notice will be sent to you prior to your expiration date; however, it is your responsibility as the Principal Investigator to timely request renewal of your approval from the Committee.

You are advised that any change in protocol for this project must be reviewed and approved by the Committee prior to implementation of the proposed change in the protocol. A protocol change/amendment form is required to be submitted for approval by the Committee. In addition, federal regulations require that the Principal Investigator promptly report, in writing any unanticipated problems or adverse events involving risks to research subjects or others.

By copy of this memorandum, the Chair of your department and/or your major professor is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols as often as needed to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Human Research Protection. The Assurance Number is IRB00000446.

Cc: Kathleen Burnett, Advisor [*****@mailer.fsu.edu]
HSC No. 2009.3805

APPENDIX E

Informed Consent Form

EXPLORING THE EFFECT OF COGNITIVE LOAD ON THE PROPENSITY FOR QUERY REFORMULATION BEHAVIOR

My name is Kyoungsik Na. I am a doctoral candidate at the Florida State University School of Library and Information Studies, working under the supervision of Dr. Kathleen Burnett. I am asking for your voluntary participation in my dissertation research. Please read the following information about the experiment I will be conducting. If you would like to participate, please sign in the appropriate box below.

Purpose of the study

The purpose of the study is to investigate the effect of cognitive load on the propensity for query reformulation behavior in information seeking on the Web (Google.com) in order to explore how cognitive load impedes the propensity of reformulating a query. This research will help better understand cognitive activities of end searchers to help system developers and developments by providing significant insight with three long-term goals:

1) To improve query formulation support and education for users;
2) To better understand searchers' query reformulation behavior; and
3) To improve search engine design and performance to better meet user needs.

If you would like to participate, you will be asked to complete the followings in consecutive order:

1) Pre-task questionnaire
2) A search task1 in Google –> Post-task questionnaire
3) A search task2 in Google –> Post-task questionnaire
4) A search task3 in Google –> Post-task questionnaire

Your search log will be recorded and it will be used for data analysis. However, the data will not be associated with your identity in anyways. The data set will be saved such as Csubj1, Csubj2, Esubj1, Esubj2, and the like. Data will be secured in a locked place and no one except the research (Kyoungsik Na) and the advisor (Dr. Kathleen Burnett) will have an access for it. All information obtained during the course of the study will remain confidential to the extent allowed by law.

Time required for participation

The estimated time for completing this experiment will approximately be an hour based on the pilot testing. Pre-survey will take less than 5 minutes and three search tasks followed by each post-survey corresponding to each search task will approximately take 50 or 60 minutes in total. Time for completing three search tasks will vary depending upon the lengthy of your time spent on the tasks.

Risks

There are no known risks to participating in this experiment, although you may become fatigued or anxious as you attempt to complete the assigned tasks. In the event that you wish to take a break or discontinue your participation, you may do so without penalty. The data will not be related to your identity in any ways to the extent allowed by law and only investigator will access it. The data will be stored in a safe place with a password protected.

If you have any questions about this research, feel free to contact:

Principal Investigator: Kyoungsik Na
Email: *****@my.fsu.edu, or
Phone: (850) 644-8117

Advisor: Dr. Kathleen Burnett
Email: *****@fsu.edu
Phone: (850) 644-8124

Human Subject Committee Contact Information:
Human Subjects Office
2010 Levy Avenue
Suite 276-C
Tallahassee, FL 32306-2743
Phone: (850) 644-7900
Fax: (850) 644-4392

Voluntary Participation

Participation in this research is completely voluntary. If you decide not to participate there will be no negative consequences. Please be aware that if you decide to participate, you may stop participating at any time and you may decide not to answer any specific question.

By signing this form I am attesting that I have read and understand the information above and I freely give my consent/assent to participate or permission to participate.

Name: ____

Signature: ____

Date: ___________/___________/_______________

REFERENCES

Aula, A. (2003). *Query formulation in web information search.* Paper presented at the Proceedings of IADIS International Conference WWW/Internet 2003.

Back, J., & Oppenheim, C. (2001). A model of cognitive load for IR: implications for user relevance feedback interaction. *Information Research, 6*(2).

Baddeley, A. (1992). Working memory. *Science, 255*, 556-559.

Baecker, R., Booth, K., Jovicic, S., McGrenere, J., & Moore, G. (2000). Reducing the gap betweenwhat users know and what they need to know. In Proceedings of theACM2000 InternationalConference on Intelligent User Interfaces (pp. 17–23). New York: ACM Press.

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern information retreival*: Addision Wesley.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407-424.

Bates, M. J. (1990). Where should the person stop and the information search interface start? *Information Processing and Management, 26*(5), 575-591.

Belkin, N. J., Cool, C., Kelly, D., Lin, S. J., Park, S. Y., Perez-Carballo, J., & Sikora, C. (2001). Interactive exploration, design and evaluation of support for query reformulation in interactive information retrieval. *Information Processing and Management, 37*, 403-434.

Bell, D., & Ruthven, I. (2004). Searchers' assessments of task complexity for web searching. *In Proceedings of the European conference on information retreival*, 57-71.

Birch, L., Juul-Kristensen, B., Jensen, C., Finsen, L., & Christensen, H. (2000). Acute response to precision, time pressure and mental demand during simulated computer work. *Scandinavian journal of work, environment & health, 26*(4), 299-305.

Brewer, J., & Hunter, A. (2005). *Foundations of multimethod research*. Thousand Oaks, CA: Sage.

Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53-61.

Bruza, P. D., & Dennis, S. (1997). Query reformulation on the Internet: Empirical data and the hyperindex search engine. *In Proceedings of the 5th RIAO conference*.

Bruza, P. D., Dennis, S., & McArthur, R. (2000). Interactive internet search: Keyword directory and query reformulation mechnisms compared. *In Proceedings of the 23rd annual ACM SIGIR conference on reserach and development in information retrieval*, 280-287.

Bystrom, K., & Jarvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing and Management, 31*(2), 191-213.

Carroll, J. M. (1995). Introduction: The Scenario Perspective on System Development. In J. M. Carroll (Ed.), *Scenario-Based Design: Envisioning work and Technology in System Development*. New York: Wiley.

Case, D. O. (2002). *Looking for information*. New York: Academic Press.

Case, D. O. (2008). *Looking for information: A survey of research on information seeking, needs, and behavior*. United Kingdom: Emerald.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition & Instruction, 8*, 293-246.

Chipperfield, B. (2006). Cognitive Load Theory and Instructional Design. Retrieved from http://www.usask.ca/education/coursework/802papers/chipperfield/index.htm

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsidertation of mental storage capacity. *Behavioral and Brain Science, 24*, 87-185.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W., & Plano-Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

Cucerzan, S., & Brill, E. (2004). Spelling correction as an iterative process that exploits the collective knowledge of web users. *Proceedings of Conference on Empirical Methods in Natural Language Processing*, 293-300.

Dennis, S., Bruza, P. D., & McArthur, R. (2002). Web searching: A process-oriented experimental study of three interactive search paradigms. *Journal of the American Society for Information Science and Technology, 53*(2), 120-133.

Dennis, S., McArthur, R., & Bruza, P. D. (1998). Searching the World Wide Web made easy? The cogntive load imposed by query refinement mechnisms. *Proceedings of the third Australian document computing symposium (ADCS'98)*, 65-71.

Efthimiadis, E. N. (1996a). Query expansion. *Annual Review of Information Systems and Technology, 31*, 121-187.

Efthimiadis, E. N. (1996b). Query expansion. In M. E. Williams (Ed.), *Annual review of information systems and technology* (Vol. 31, pp. 121-187). Medford, NJ: Information Today.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211-245.

Field, H., Allan, J., & Jones, R. (2010). Predicting searcher frustration. In SIGIR '10: Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval, pp. 34-41, doi:10.1145/1835449.1835458

Fiske, S. T., Gilbert, D. T., & Lindzey, G. (Eds.). (2010). *Handbook of Social Psychology* (5th ed.): Wiley.

Galy E., Cariou, M., & Melan, C. (2011). What is the relationship between mental workload factors and cognitive load types? *Internationial Journal of Psychophysiology*, doi:10.1016/j.ijpsycho.2011.09.023

Goodwin, C. J. (2009). *Research in Psychology: Methods and Design*: Wiley.

Grossman, D. A., & Frieder, O. (2005). *Information retrieval: algorithms and heuristics*: Kluwer Academic Pub.

Gu, Y. & Gilbert, J. E. (2004). The development of a Voice System that interacts with astudent information database. In Proceedings of the 42nd Annual Southeast Regional Conference (Huntsville, Alabama, April 02 - 03, 2004). ACM-SE 42. ACM, New York, NY, 248-252.

Gwizdka, J. (2008). Cognitive load on Web search tasks. *In Proceedings of Workshop on Cognition and the Web*, 83-86.

Gwizdka, J. (2010). Distribution of cognitive load in web search. *Journal of American Society for Information Science and Technology, 61*(11), 2167-2187.

Haythornthwaite, C. (2001). The internet in everyday life. *American Behavioral Scientist, 45*(3), 363-382.

Hearst, M. A. (2009). *Search user interfaces*. New York, NY: Cambridge University Press.

Hsieh-Yee, I. (1993). Effects of Search Experience and Subject Knowledge on the Search Tactics of Novice and Experienced Searchers. *Journal of American Society for Information Science, 44*(3), 161-174.

Hu, P. J. H., Ma, P. C., & Chau, P. Y. K. (1999). Evaluation of user interface designs for information retrieval systems: a computer-based experiment. *Decision Support Systems, 27*(1-2), 125-143.

Ingwersen, P. (1992). *Information retrieval interaction*. London: Taylor Graham.

Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of Documentation, 52*(1), 3-50.

Ingwersen, P. (2000). Cognitive information retrieval. In M. E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 34, pp. 3-51). Medford, NJ: Information Today, Inc.

Jansen, B. J. (1998). The Graphical User Interface: An Introductini. *SIGCHI Bulletin. 30*(2), 22-26.

Jansen, B. J. (2005). Seeking and implementing automated assistance during the search process. *Information Processing & Management, 41*(4), 909-928.

Jansen, B. J. (2006a). Search log analysis: What is it; what's been done; how to do it. *Library & Information Science Research, 28*(3), 407-432.

Jansen, B. J. (2006b). Using temporal patterns of interactions to design effective automated searching assistance systems. *Communications of the ACM, 49*(4), 72-74.

Jansen, B. J., & McNeese, M. D. (2005). Evaluating the effectiveness of and patterns of interactions with automated searching assistance. *Journal of the American Society for Information Science and Technology, 56*(14), 1480-1503.

Jansen, B. J., & Spink, A. (2005a). An analysis of Web searching by European Alltheweb.com users. *Information Processing & Management, 41*(2), 361-381.

Jansen, B. J., & Spink, A. (2005b). How are we searching the World Wide Web? A comparison of nine search engine transaction logs *Information Processing & Management, 42*(1), 248-263.

Jansen, B. J., Spink, A., Blakely, C., & Koshman, S. (2006). Web searcher interactions with the Dogpile.com meta-search engine. *Journal of the American Society for Information Science and Technology, 58*(4), 1875-1887.

Jansen, B. J., Spink, A., & Koshman, S. (2007). Web searcher interaction with the Dogpile.com meta-search engine. *Journal of the American Society for Information Science and Technology, 57*(5), 744-755.

Jansen, B. J., Spink, A., & Pedersen, J. (2005). Trend analysis of AltaVista Web searching. *Journal of the American Society for Information Science and Technology, 56*(6), 559-570.

Jansen, B. J., Spink, A., & Saracevic, T. (2000). Real life, real users and real needs: A study and analysis of users' queries on the Web. *Information Processing and Management, 36*(2), 207-227.

Jansen, B. J., Zhang, M., & Spink, A. (2007). Patterns and transitions of query reformulation during web searching. *International Journal of Web Information Systems, 3*(4), 328-340.

Jones, K. S. (1981). *Information retrieval experiment*. London: Butterworths.

Kim, Y. M., & Rieh, S. Y. (2005). *Dual-task performance as a measure for mental effort in library searching and Web searching.* Paper presented at the In Proceedings of the 68th Annual Meeting of ASIS&T.

Klink, S. (2001a). Query reformulation with collaborative concept-based expansion. *Proceedings of the First International Workshop on Web Document Analysis* 19-22.

Klink, S. (2001b). Query reformulation with collaborative concept-based expansion. *In Proceedings of the First International Workshop on Web Document Analysis*.

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.

Mandel, T. (2002). *The elements of user interface design*. New York, NY, USA: John Wiley & Sons, Inc.

Manning, C. D., Raghavan, P., & Schutze, H. (2008). *Introduction to information retrieval*: Cambridge University Press.

Marchionini, G. (1995). *Information seeking in electronic environments*. New York: Cambridge University Press.

Marchionini, G., & White, R. (2007). Find what you need, understand what you find. *International Journal of Human-Computer Interaction, 23*(3), 205-237.

Marchoionini, G., & Shneiderman, B. (1988). Finding facts vs. browsing knowledge in hypertext systems. *IEEE computer, 21*(1), 70-80.

Matthews, R, Legg, S, & Charlton, S. (2003) The effect of cell phone type on drivers subjective workload during concurrent driving and conversing. *Accident Analysis and Prevention, 35*, 451-457.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge, United Kingdom: Cambridge University Press.

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

Moreno, R., & Park, B. (2010). Cognitive load theory: Historical development and relation to other theories. In J. L. Plass, R. Moreno & R. Brunken (Eds.), *Cogntive load theory*. Cambridge, New York: Cambridge University Press.

Noyes, J. M., & Bruneau, D. P. J. (2007). A self-analysis of the NASA-TLX workload measure *Ergonomics, 50*(4), 514-519.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In Boff, K., Kaufman, L., & Thomas, J. (Eds.), Handbook of Perception and Human Performance, Vol. II (pp. 42-1 to 42-49). New York: Wiley.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1-4.

Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63-72.

Paas, F., & Van Merrienboer, J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive load approach. *Journal of Educational Psychology, 86*, 122-133.

Pew Internet & American Life Project tracking surveys. (2008). Retrieved from http://www.pewinternet.org/Trend-Data/Online-Activites-Total.aspx

Pew Internet & American Life Project tracking surveys. (2011). Retreived from http://www.pewinternet.org/Reports/2011/Search-and-email.aspx

Pew Internet & American Life Project tracking surveys. (2011). Retreived from http://www.pewinternet.org/Reports/2011/Cell-Phones.aspx

Pirolli, P. (2007). *Information Foraging Theory: Adaptive Interaction with Information*: Oxford University Press, USA.

Pirolli, P., & Card, S. K. (1999). Information foraging. *Psychological Review, 106*, 643-675.

Raghavan, V., & Jung, W. (1989). A critical investigation of recall and precision as measures of retrieval system performance. *ACM Transactions on Information Systems, 7*(3), 205-229. .

Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: Effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology: Applied, 9*(2), 119.

Rieh, S. Y., & Xie, H. (2001). Patterns and sequences of multiple query reformulations in Web searching: a preliminary study. *Proceedings of the 64th annual meeting of the American Society for Information Science and Technology, 38*, 246-255.

Rieh, S. Y., & Xie, H. (2006). Analysis of multiple query reformulations on the Web: the interactive information retrieval context. *Information Processing & Management, 42*(3), 751-768.

Rosson, M. B., & Carroll, J. M. (2002). *Usability Engineering: Scenario-based Development of Human-Computer Interaction*. London: Academic Press.

Salton, G., & Buckley, C. (1990). Improving retrieval performance by relevance feedback. *Journal of the American Society for Information Science, 41*(4), 288-297.

Saracevic, T. (1996). Relevance reconsidered. Information science: Integration in perspectives. *Proceedings of the Second International Conference on Conception of Library & Information Science (CoLIS 2)*, 201-218.

Saracevic, T. (2007a). A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology, 58*(3), 1915-1933.

Saracevic, T. (2007b). A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance. *Journal of the American Society for Information Science and Technology, 58*(13), 2126-2144.

Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A re-examination of relevance: toward a dynamic, situational definition. *Information Processing & Management, 26*(6), 734-744.

Schmutz, P., Heinz, S., Metrailler, Y., & Opwis, K. (2009). Cognitive load in eCommerce application - measurement and effects on user satisfaction. *Advances in Human-Computer Interaction, 2009*, 1-9.

Schneiderman, B., & Plaisant, C. (1998). *Designing the user interface*: Addison-Wesley Longman.

Shafi, S. M., & Rather, R. A. (2005). Precision and Recall of Five Search Engines for Retrieval of Scholarly Information in the Field of Biotechnology. *Webology, 2*(2). Retrieved from http://www.webology.org/2005/v2n2/a12.html

Shneiderman, S. B., & Plaisant, C. (2005). *Designing the user interface 4 th edition*: Pearson Addison Wesley, USA.

Sihvinen, A., & Vakkari, P. (2004). Subject knowledge improves interactive query expansion assisted by a thesaurus. *Journal of Documentation, 60*(6), 673-690.

Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1999). Analysis of a very large Web search engine query log. *SIGIR Forum, 33*(1), 6-12.

Spector, J. M., Merrill, M. D., Merrienboer, J. V., & Driscoll, M. P. (Eds.). (2007). *Handbook of Research on Educational Communications and Technology*: Routledge.

Spink, A., Jansen, B. J., & Ozmultu, C. (2001). Use of query reformulation and relevance feedback by Excite users. *Internet Research, 10*(4), 317-328.

Spink, A., Wolfram, D., Jansen, B. J., & Saracevic, T. (2001). Searching the Web: the public and their queries. *Journal of the American Society for Information Science, 52*(3), 226-234.

Stojanovic, N. (2005). On the conceptualization of the query refinement task. *Library Management, 26*(4/5), 231-293.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction, 4*, 295-312.

Sweller, J. (1999). *Instructional design in technical areas*. Melbourne: ACER Press.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology: General, 119*, 176-192.

Sweller, J., Van Merrienboer, J., & Pass, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Toffler, A. (1970). *Future shock*. New York: Bantam Books.

Vakkari, P. (1999). Task complexity, problem structure and information actions, integrated studies on information seeking and retrieval. *Information Processing & Management, 35*, 819-837.

Vakkari, P. (2000). *Cognition and changes of search terms and tactics during tack performance: a longitudinal study*. Paper presented at the Proceedings of the RIAO' 2000 conference, Paris: C.I.D.

Vakkari, P., Pennanen, M., & Serola, S. (2003). Changes in search terms and tactics while writing a research proposal: a longitudinal case study. *Information Processing & Management, 39*, 445-463.

Voorhees, E. M., & Harman, D. K. (Eds.). (2005). *TREC: Experiment and Evaluation in Information Retrieval*. Cambridge, MA: The MIT Press.

White, R., & Marchionini, G. (2007). Examining the effectiveness of real-time query expansion. *Information Processing & Management, 43*, 685-704.

White, R. W., & Roth, R. A. (2009). *Exploratory search: Beyond the query-response paradigm*: Morgan & Claypool Publisher.

Wikipedia. (2010). Google Retrieved March, 20, 2010, from http://en.wikipedia.org/wiki/Google

Wickens, C. D. (1992). Engineering psychology and human performance (2nd Edition). New York: Harper Collins.

Wilkinson, S. C., Reader, W., & Payne, S. J. (2012). Adaptive browsing: Sensitivity to time pressure and task difficulty. *International Journal of Human-Computer Studies, 70*(1), 14025.

Xie, B., & Salvendy, G. (2000). Predition of mental workload in single and multiple task environments. *International Journal of Cognitive Ergonomics, 4*(3), 213-242.

Yurko, Y. Y., Scerbo, M. W., Prabhu, A. S., Acker, C. E., & Stefanidis, D. (2010). Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. *Simulation in Healthcare, 5*(5), 267.

Zazo, A. F., Figuerola, C. G., Berrocal, J. L. A., & Rodriguez, E. (2005). Reformulation of queries using similarity thesauri. *Information Processing & Management, 41*, 1163-1173.

BIOGRAPHICAL SKETCH

Kyoungsik Na

School of Library & Information Studies
College of Communication & Information
Florida State University
101 Louis Shores Building
Tallahassee, FL 32306-2100

Education

Ph. D. Candidate, Library & Information Studies, Florida State University, Tallahassee

Completed:	April, 2012
Major:	Information Studies
Minor:	Instructional Technology System and Design
Supervisory committee:	Dr. Kathleen Burnett (Chair), Dr. Gary Burnett, Dr. Besiki Stvilia, Dr. Neil Charness (Cognitive Psychology)
Dissertation title:	"*Exploring the Effect of Cognitive Load on the Propensity for Query Reformulation Behavior*" This exploratory study is to examine the effect of cognitive load on the propensity to reformulate queries during information seeking on the Web, specifically the effect of manipulating three affective components of cognitive load -- mental demand, temporal demand, and frustration -- is explored to in relation to the propensity to reformulate queries.

M. S. Information Science, University of Tennessee, Knoxville, Fall 2005

B. S. Computer Science, University of Tennessee, Knoxville, Fall 2002