FACILITATING BROWSING WITH INFORMATION VISUALIZATION: IS ANIMATION A POWERFUL SCENT?

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

By
Stella Taylor
May, 2009

ABSTRACT

Title: Facilitating Browsing with Information Visualization: Is Animation a Powerful Scent?
Candidate's Name: Stella Taylor
Degree: Philosophy of Degree
Temple University, 2009
Doctoral Advisory Committee Chair: David Schuff

Search engines make vast amounts of information available to Internet users. Two types of tasks users engage in using search engines are closed-ended and open-ended. For closed-ended tasks, individuals have narrow objectives that require finding specific results. For open-ended tasks, individuals only have general objectives that require finding as much relevant information as possible about a topic, which can be difficult when large numbers of both relevant and irrelevant results are returned from a query. This can also leave users in a state of information overload. Some search engines have incorporated information visualization techniques (combining cognitive senses with visual cues that allow for better understanding the information) to facilitate browsing through results in order to reduce information overload. However, there is little research that identifies which visual cues are the most desirable for the presentation of search results.

According to information foraging theory, cues that have strong scents will help users find information faster. In this study, we investigate the effects of augmenting visualizations with animation as a powerful scent to help users more easily identify

relevant information in search engine results. This study employs cognitive fit theory to study the effect of different information formats on users' performance in completing the two different tasks.

Overall, we find evidence that the effectiveness of cues such as animation is task-dependent. For example, we find that visualizations with animation are less effective than a standard textual display for subjects performing closed-ended web search tasks. The results of this study have strong implications for integrating appropriate cues into visualizations in order to help people find information.

ACKNOWLEDGMENTS

When I was first notified that I was accepted into the doctoral program, I was so excited and overjoyed. However the journey to get my doctorate was not always filled with joy. The road involved patience, strength, maturity through experience and sometimes tears. Throughout my journey, there have been many individuals that have helped me along the way.

I would first like to thank Dr. David Schuff, my doctoral advisor and committee chair, for his understanding, guidance and patience. Dr. Schuff was instrumental in the success of this study. He was very encouraging during my years in the program by developing my research skills. Dr. Schuff inspired me to reach new levels of thinking as a researcher. His insightful comments and guidance during all stages of this research were most helpful and important to the progress of this study. I am also grateful for Dr. Schuff's immediate response to the numerous emails and phone calls that I made during my research studies. Dr. Schuff has helped me overcome many obstacles during this research study and I appreciate all his support. Thank you very much, Dr. Schuff.

I am very grateful for my committee members, Dr. Ozgur Turetken, Dr. Youngin Yoo and Dr. Dennis Galletta, for their continual support and encouragement. Their guidance has helped me extend my research focus and methodology. Through his vast knowledge in visualization, Dr. Turetken helped further the focus of this study. Dr. Turetken was very helpful with the numerous statistical questions that I had. Dr. Yoo always pushed me to higher standards of research through his vast amount of knowledge and experience. I appreciate their insightful comments and edits to my paper. I extend my true gratitude to the wonderful committee that supported and guided me through this long process. Dr. Galletta is truly an amazing individual and researcher. Throughout my research, he provided me insight and papers regarding cognitive fit theory. He never hesitated to provide me information and advice. He went out of his way to help me during the most critical time of this process. I am truly indebted to Dr. Galletta.

I am truly grateful to Dr. Munir Mandviwalla, who has always been supportive and encouraging throughout my studies. He provided sound advice and graciously helped with any “bumps” that came across my path. He always provided a “lending ear”, a few pieces of chocolate, or a laugh through his “odd” sense of humor.

I would like to thank Dr. Heinz Klein, for his support and guidance during my early years in the program. His knowledge in qualitative research helped me develop my research skills.

I truly appreciate the work that Juhi Ghosh did in creating the prototypes used in this study. She was very supportive during the design phase of this study. Even after she graduated, Juhi would still respond and help troubleshoot the technical issues that would arise. I truly appreciate all the support that Juhi provided.

I would like to extend my gratitude to the numerous professors that have helped me throughout my doctorate program: Dr. Ravi Patnayakuni, Dr. Dan Fesenmaier, Dr. Joseph DuCette, Dr. Sunil Wattal, and Dr. Paul Pavlou. I extend many thanks to Dr. Mary Conran who helped me recruit students within the Marketing Department for participation in my experiments. I am also grateful for the support that Institute for Business and Information Technology (IBIT) provided.

I also would like to extend my gratitude to my management and my co-workers in the DDPM group at Accenture for their continual support to further my education.

I appreciate all the individuals that participated in the experiments for this study. The numerous friends, the employees at Accenture, and the students at Temple University were enthusiastic in participating in my study and were helpful in the success of this study.

I would also like to thank the following individuals who helped with the data collection of this study:

Joseph Allegra
Dr. Zaheeruddin Asif
Craig Atwater
Marlene Bell
Dr. Anthony DiBenedetto
Matthew Dampf
Mart Doyle
Dr. Ram Gopalan
Brian Healy
Joyti Johal
Jeannette-Marie Kelley
Suzanne Chehayeb Makarem
Cindy Joy Marselis
Damon McNeil
Dr. Susan Mudambi
Julie Phillips
Doug Schultz
Dr. Harry Shrank
Dr. Paul Weinberg
Dr. Dan Zhang

Most importantly, this would have never been done without the love and patience from my family and friends. I am thankful for my many friends, who allowed me to vent about theories, qualitative and quantitative analysis, statistical approaches, and many other research issues even though they had no idea what I was talking about. I am truly grateful for the support that Julie provided during the roughest moments of my research progress.

I am most grateful to my wonderful family, whose support was immeasurable. I am thankful for their continuous prayers. I am grateful for my niece and nephews for being able to always put a smile on my face. I extend a special thanks to two angels that God sent me here on Earth: my Mom and Dad. My parents came to this country so that my siblings and I could experience the best type of education and reach for goals that would not have been possible in India. I am proud to dedicate this dissertation to my parents, whom have been a constant source of love, concern, support, and strength throughout these years. By example, they have shown me patience, perseverance, faithfulness and vigor. My parents have always instilled confidence in me and have always encouraged me through my studies even though they didn’t understand all the aspects of my research. I do not have enough words to express how grateful I am to my parents.

Last, I would like to thank Almighty God. None of this would be possible if it was not the strength of the Lord by my side. No matter how difficult the journey, I can do all things through Christ.

“For I know the plans I have for you," declares the LORD,

"plans to prosper you and not to harm you,

plans to give you hope and a future.”

JEREMIAH 29:11

I dedicate this dissertation to my parents, to whom I owe everything I am today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

"What information consumes is rather obvious: It consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention"

– Herbert Simon

1.1 Problem Statement

Can you imagine not having a mechanism to search for information? In 1993, it was documented that about 130 websites were available on the Internet and there seemed to be no need for search engines at that time. WebCrawler (www.webcrawler.com), which was the first search engine that provided full text search, was created in 1994 and Google was developed in 1998. Now in 2009, Google is such a popular search engine that the verb "google," was added to the Merriam Webster Collegiate Dictionary and the Oxford English Dictionary in 2006, meaning, "to use the Google search engine to obtain information on the Internet." Can you imagine life now in 2008 without a search engine on the Internet? Imagine trying to find articles on global warming or trying to find an Italian restaurant in the area.

Bawden (2001) found that the number of information sources is rising especially quickly on the Internet. In March 2008, it was documented that there are about 162 million web pages on the Internet (http://royal.pingdom.com/?p=273). According to Pew Internet & American Life Project Tracking surveys (2004), one of the most frequent activities on the Internet is the use of search engines. In addition to the rise of information sources on the Internet, there is simply an abundant amount of electronic information available to individuals in the workplace or school. When the amount of available information exceeds an individual's ability to process it is termed "information overload" (Eppler and Mengis, 2004; Schick et al., 1990)

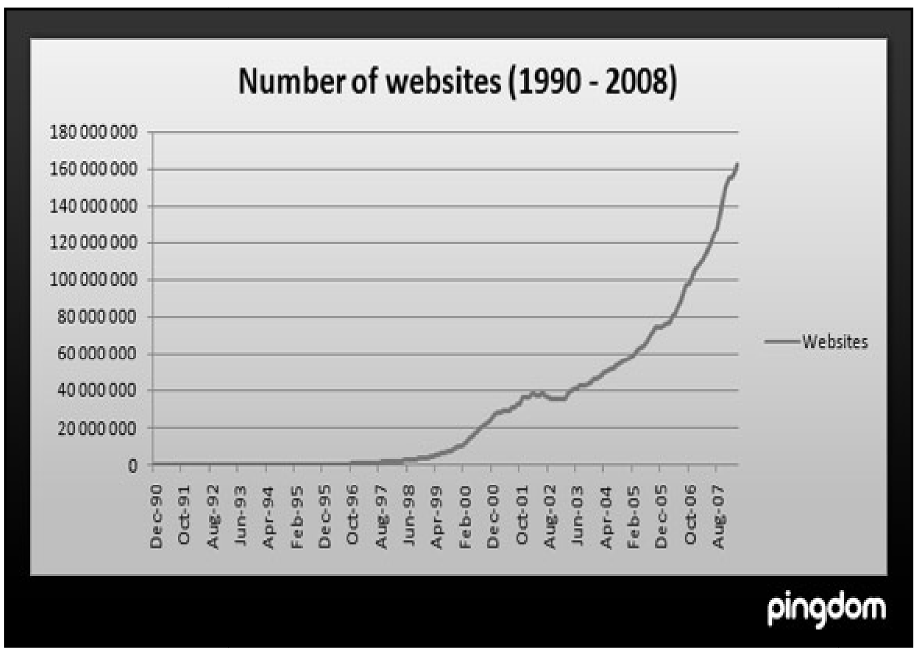


Figure 1. Number of Websites (1990-2008)

Sites = # of web servers (one host may have multiple sites by using different domains or port numbers) (extracted from http://royal.pingdom.com/?p=273)

Individuals tend to become frustrated with information overload when browsing the Web. Research has also shown that when completing tasks on the Internet, individuals stop browsing after 20 to 30 results from a search engine, whether or not they have completed their search (Roussinov and Chen, 2001). Two types of web search tasks individuals perform are closed-ended searches for specific information and open-ended searches for general topics. Closed-ended tasks involve a specific objective and finding accurate answers. For example, an MSN Live or Google search for the query "who won the Noble Prize in Medicine in 1979?" would produce a specific answer. In closed-ended tasks, individuals tend to find results rather quickly since the top results from a query on a search engine are likely to have the answers. Open-ended tasks involve finding relevant information on a general topic. Individuals may have to browse many search results to fully understand the topic and make a valid judgment to complete the task. Broad queries used in open-ended browsing tasks can produce numerous results, both relevant and irrelevant, which can be a problem when trying to complete a task. For this reason, open-ended tasks are difficult to complete compared to more solvable closed-ended tasks. In fact, Chung et al. (2005) claim that "search engine displays often overwhelm users with irrelevant information" (p. 58). Search engine displays range from a standard ranked list to a presentation of results through different cues. In this study, we utilize cognitive fit theory to identify which tasks match with various formats.

Information visualization is a collection of techniques for combining cognitive senses with visual cues that allow for better understanding of the information (Turetken and Sharda, 2004). The presentation of search engine results is a potential application of information visualization, which can reduce information overload by shifting some of the information processing load to sensory systems. Individuals are able to more easily understand information when it is presented (Tufte, 2001) in a graphical format. The use of visualization techniques should allow individuals to more easily navigate through search results by directing the user to the most relevant results by providing the user with a combination of visual cues and an understandable structure.

However, even with visualization, finding relevant information is a difficult task (Dumais et al., 2001; Roussinov and Chen, 2001; Turetken and Sharda, 2005) since there is a lack of appropriate cues available to help users navigate through the large information space that a search creates. According to Dumais et al. (2001), there is little research that identifies which cues are the most desirable in the visual presentation of search results. Card et al. (1999) suggest that other features, such as animation, have been underutilized in visualization. In an analysis of web space visualization, Turetken and Sharda (2007) state “a great majority of the systems surveyed…do not use animations in spite of the technical feasibility of these visual aids” (p. 77). Drawing from information foraging theory, motion effect theories, and the construct of distinctiveness; we propose that animation can be a useful information visualization technique since animation can draw users’ attention to the most relevant information. In this study, we prototype a visual format that incorporates animation to help individuals navigate search engine results. Based on information foraging theory and the results of this study, we are able to identify the strength of different scents utilized in the designed formats and learn if animation can be a powerful scent.

Therefore, we investigate the following research question:

What is the effect of animation in facilitating higher-information seeking performance in open-ended and closed-ended tasks?

1.2 Study Goals and Objectives

In addition to gaining insight regarding the usefulness of animation in the presentations of search results, there is also a need to identify for which type of visual format and browsing tasks, animation is most suitable. Accordingly, a second goal of this study is to identify which type of visual format is most suitable for open-ended versus closed-ended browsing tasks. We develop several formats to display search results, and using cognitive fit theory as a theoretical framework, we propose a research model to explain how matching format to task (closed-ended versus open-ended) reduces the cognitive effort required to complete the task. By investigating different visualization formats and tasks, we can identify whether animation is a useful cue to achieve higher task performance and user satisfaction.

The contributions of this research are twofold. From a research perspective, it integrates theories from the research domains of information systems and advertising to help understand the effects of animation on the visualization of search results. We will also gain insight regarding how the effects of animation vary across different task types. From a practical perspective, the findings will help web designers determine how to best present search results to end users. We will understand the usefulness of animation as a cue in the visualization of search engine results.

1.3 Organization of Dissertation

This paper is organized in ten chapters. Chapter 2 begins with the research rationale for this study. In chapter 3, previous literature on visualization is discussed. The theoretical framework for this study is discussed in Chapter 4 and the proposed hypotheses based on the theories are provided in Chapter 5. In Chapter 6, the steps in designing the prototypes or formats utilized in this study are described. Chapter 7 provides the research methodology that this study followed and Chapter 8 provides the statistical analyses of the data collected. Discussions of the results found in this study are provided in Chapter 9 and finally, Chapter 10 discussed the conclusions and future work to enhance this field of study.

CHAPTER 2

RESEARCH RATIONAL

2.0 The Problem: Information Overload

Information overload is the state when individuals are not able to process the information available to them. This problem has been studied in different areas such as accounting (Schick et al., 1990), marketing (Keller and Staelin, 1987) and information systems (Schultze and Vandenbosch, 1998). Although information overload is seen in many areas, the Internet has become a major contributor to this phenomenon (Swash, 1998) because of its sheer volume of content. An overabundance of irrelevant information tends to frustrate individuals (Ackoff, 1967). Schneider (1987) found that information generally could be ambiguous and complex. Users become uncertain of identifying relevant information among extraneous information, which is often the problem with open-ended tasks.

Previous research has focused on the effect of overload on an individual's decision-making performance (Eppler et al., 2004). The burden of information overload confuses individuals and affects their ability to set priorities, make decisions, and recall prior information (Schick et al., 1990). This affects decision making in two ways: users are not able to locate what they need and they often overlook critical information. In open-ended tasks, it is important for individuals to find relevant information and not to be confused with irrelevant information.

Ho and Tang (2001) state that three factors cause information overload: information quantity, information quality, and information format. Most of the issues with search results refer to tasks completed using ranked textual list, a common presentation format provided by traditional search engines. As stated earlier, open-ended tasks completed on textual formats produces numerous results. The objective of open-ended and closed-ended tasks is to find relevant information and results. Other formats besides the standard textual list need to be identified to discover whether certain tasks can be completed while reducing information overload.

2.1 The Solutions: Reducing Information Overload

Many researchers have proposed frameworks and techniques to deal with information overload. Shneiderman (1996) proposed a task by data type taxonomy (TTT) to study the types of data and tasks involved in visual displays of textual information. One such display is linear which includes a result-list of hyperlinks. Result lists are a one-dimensional data type still widely used by many search engines due to its simplicity in presenting results. However, as discussed earlier, result lists only allow limited browsing since users have to scroll through many pages to see all the results. Other data types such as two-dimensional data, tree data, and network data allow for more effective browsing to be performed. These data types also support visual capabilities, which can reduce information overload as explained through the Model Human Processor (MHP).

MHP is a psychological model that shows how sensory buffers, short-term and long-term memory, interact with information from senses to produce responses to information-related tasks (Card et al., 1983). Sensory buffers are stores for stimuli received by the senses (visual and auditory). Unless encoded in the short-term memory, this information is quickly lost. Short-term memory (working memory) acts as a store for information that is required quickly. Long-term memory is the main source for memory. In order for individuals to complete tasks, information is taken from short-term and long-term memory. Clustering of information is one way that enables individuals to utilize visual senses. Therefore, overload is reduced since some of the processing is shifted to the sensory systems. (Card et al., 1983; Turteken and Sharda, 2007)

Researchers have investigated how the use of filtering and clustering algorithms can help reduce information overload. While filtering allows users to cope with large amounts of information, it does not eliminate the problem completely since the number of results can still be too large to process, resulting in users still overlooking relevant information. In addition to filtering, clustering is also a common technique used to help with overload as it provides structure by grouping similar results together. Previous studies (Cutting et al., 1992; Hearst, 1995; Turetken and Sharda, 2005) have demonstrated to usefulness of clustering in helping individuals easily recognize which groups of results are of interest and eliminate groups that are not needed, thereby reducing information overload.

In addition to clustering, visualization allows individuals to increase sensory buffers (visual and auditory senses) so that short-term memory is increased. There are many additional benefits that visualization can provide. Card et al. (1999) suggest six major ways that visualization can increase cognition by: 1) increasing memory and processing resources for individuals, 2) reducing the search for information, 3) using visual presentations to enhance recognition of relationships, 4) making complex problems visually simple, 5) using cues to provide mechanisms for attention, and 6) providing information in a manner that can be manipulated.

Turetken and Sharda (2004) state “the relative processing capacity and speed advantage of the perceptual (visual) system to the cognitive system results in the better and quicker understanding of information when supported by visual cues” (p. 416). Visualization reduces the search of data by grouping related information together (Card et al., 1999). Patterns can also be detected and new knowledge can be discovered through finding new relationships. In the next section, previous studies that utilize visualization and various techniques are discussed.

2.2 The Answer: Visualization

A number of studies report that insight and problem-solving performance can be improved with appropriate visualizations (Crapo et al., 2000; Pinker, 1997; Hong and O’Neil, 1992).Wise et al. (1995) developed an application that transformed text content into a spatial representation that enabled enhanced visual browsing and analysis. They stated that the application enabled users to avoid language processing, thereby reducing mental workload. Their contention was that this creates “an interaction with text that more nearly resembles perception and action with the natural world than with the abstractions of written language” (p. 51).

Turetken and Sharda (2005) developed a fisheye interface of a zoomable two-dimensional map of clustered search results called FISPA. Each cluster was labeled by the most frequent term in the cluster. The zooming feature allowed users to interact with the results and filter on specific topic areas so that users can view all the results in that category at the same time. When a user zoomed in on a cluster or category and found an individual result, a separate window displayed the target web page. They found that subjects completed search tasks faster with their visual interface than a textual interface. Wise et al. (1995) developed two different visualization approaches to the same information in a document repository. The Galaxies visualization clusters similar documents together into a 2D scatterplot of "docupoints." Users were able to understand patterns and trends in the documents, and were able to explore each cluster to identify important documents within it. Users stated that they had "enhanced insight and time savings such as 'discovering in 35 minutes what would have taken two weeks otherwise'" (Wise et al., 1995, p. 56).

Zamir and Etzioni (1999) developed an interface called Grouper, which presented the results from a meta-search engine, Husky Search. The interface grouped the search results into clusters labeled by phrases. The results of the interface were presented in a ranked list with a summary of a cluster and a percentage of the documents within the cluster.

Allan et al. (2001) developed the Lighthouse system, which presents results from search engines visually through document similarities. The visualization is created through an algorithm where similarities between documents are represented as a physical distance between the documents within the display. They found that users were more successful using the visualization format over the textual format. Next, we will discuss different visualization formats.

Chung et al. (2005) developed two browsing methods: a knowledge map and web community. Web community uses a genetic algorithm to organize Web sites into a tree format, and the knowledge map uses a multidimensional scaling algorithm to place Web sites as points on a screen. Their study compared a ranked result list to the search result display from the search site Kartoo (http://www.kartoo.com) in addition to the two browsers the authors developed. Kartoo displays search results in a map format, with circles representing Web sites and lines linking the Web sites with similar key words. Subjects were asked to perform open-ended and closed-ended tasks regarding specific topics, using the various formats to find accurate and relevant results. They found that web communities were better than ranked lists in terms of all performance measures (effectiveness, efficiency and usability). They stated that “appropriate use of visualization was the main contributor to superior performance” (p. 81). The knowledge map was created using MDS to display a two-dimensional representation of the search results. Chung et al. (2005) found that knowledge maps were significantly better than textual lists across performance measures (effectiveness, efficiency and usability). The authors stated “clustering and visualization were the main contributing factors” to this result (p. 76). However, since the knowledge map did not have a zooming capability, subjects rated the knowledge map less usable. For usability, subjects rated web community significantly higher than ranked result lists due to visualization effects, clustering, labeling, and providing details on demand. One subject stated “Once I spot the label, I can move to the relevant topics very easily….(web community save(s) time, (I) don’t need to read all the summaries and Web pages to decide which are relevant" (p 77). Fifteen other subjects had similar comments. With regard to visualization effects, one subject stated that "visualization helps to navigate faster and easier" (p. 77). The authors believe that clustering and visualization of web community contributed to its higher rating in the study.

Based on previous studies utilizing visualization, performance has shown to improve. In this study, we will investigate how visualization formats will help users find relevant results more quickly for certain web tasks.

2.3 Summary

This chapter reviews the research rationale for using visualization to reduce information overload and enable successful performance. Techniques and cues are useful in visualization which will be described in the next chapter. We will also review the process of visualization and describe literature categorized by the different techniques or cues most relevant to this study, which are map formats, clustering, color, size and animation.

CHAPTER 3

REVIEW OF VISUALIZATION

As discussed earlier, visualization can reduce information overload since images stimulate the sensory buffers. The goal of visualization is to provide insight into the information contained within one or more documents without having to read those documents (Wise et al., 1995). This is an important feature for system designers. We will begin by discussing the process of clustering and visualizing a set of documents. Visualization consists of three stages: document analysis, the application of algorithms to organize those documents, and the visualization of the resulting groupings (Spence, 2001). We describe each stage below.

3.1 Visualization Process

3.1.1 Document Analysis

Document analysis is described as "extracts the essential descriptors of a collection of text, usually according to the interests of a user expressed as a set of key words" (Spence, 2001). Document analysis can provide users with numerous results based on their query. On the Web, a technique used in document analysis is web-mining. Three phases of web mining are web content, web structure, and web usage.

Web-content mining treats web pages as content keywords that can be accessed. Web-structure mining uses a network model of the Web to determine the importance of the web page. For example, the number of links that point to that web page determines that page's importance. Web-usage mining performs data mining on web logs. Web pages that are associated with a web server are analyzed to determine their importance. This enables a search engine to provide multiple results that might be relevant to a query based on keywords within the page or the importance of that page. For example, Google and MSN Live Search utilize web mining to place results relevant to a query higher on the result page.

3.1.2 Algorithm Application

Appropriate clustering algorithms can be created to provide an efficient representation of the documents retrieved through analysis. Algorithms have been used to cluster results and project the results into two- or three-dimensional space. Clustering algorithms classify objects into meaningful subsets of closely related documents. Clustering attempts to group documents together based on similarities of selected attributes. Therefore, documents relating to a specific topic would be placed together in one cluster, provided the attribute selected is content. In this study, we will use clustering algorithms to categorize groups of search results.

Clustering web search results allow individuals to find documents more easily as well as get an overview of the group of documents. Clustering can be performed on documents in advance of or after retrieval. However, performing clustering after retrieval has showed the best results (Zamir and Etzioni, 1999) because it is based on the documents obtained from queries. If clustering is done prior to retrieval, then documents that are not related to the query will be clustered and influence the cluster algorithm, resulting in poorly formed clusters.

Zamir and Etzioni (1999) enumerated several key requirements for post-retrieval document clustering: coherent clusters, efficient browsing and speed. They found that documents that have multiple topics should not be placed into one cluster. Therefore, the algorithm should have the ability to overlap. Clustering should enable individuals to identify groups of results easily through labels. The clustering algorithm should be able to cluster thousands of documents easily in a few seconds. Since clustering is proven to be useful for users, this study will cluster similar query results into categories to provide better filtering of finding relevant information.

3.1.3 Presentation

Visualization is the process of displaying data in a visual format. Lin (1997) proposes four types of visual formats for data: hierarchical displays, network displays, scatter displays, and map displays. Hierarchical displays show the data in a hierarchical form, which simplifies complex data structures. Hierarchical displays separate data through different levels, branches, and clusters (Lin, 1997). This type of visual format provides a global and local view of data as well as provides individuals with the ability to direct attention to a particular level or branch. Lin (1997) states disadvantages of the hierarchical display are oversimplification of structures and increased cognitive load for users.

The second format Lin (1997) discusses is the network display. A network display provides a graphical display of links and nodes. Network displays show structures on the screen and allow individuals to follow the link to browse the nodes (documents). These types of displays can provide a more general and complex structure than hierarchical display. However, some disadvantages of network display are that complex structures can sometimes confuse and distract users. It is also difficult to show the global view of the structure since the display space is limited.

Scatter displays represent graphical dotted images of data in a two-dimensional visual space (Lin, 1997). This display shows dots or other small icons that represent individual data points. Scatter displays are generated through algorithms and are most useful for statistical data. The disadvantage of scatter displays is that there is a slight distortion due to reduction of all the data points into a two-dimensional representation, which results in a slight error in the degree of similarity among results (or data points). In addition, in scatter displays, there is no path for following links as there are in hierarchical and network displays.

Another visual format is the map display. According to Lin (1997), the map display is the "best example of using graphical displays to show large amounts of information and their relationships" (p. 44). The map is a display of the document space rather than a representation of the data. This type of format provides rich visual information and provides information in different levels to allow individuals to interact and find relationships with the data. To ensure that map displays are useful, Lin (1997) states that maps should be simple to understand and support visual associations. In this study, our visualization format will be a map display.

Map displays provide rich visual information with the ability to form relationships with the data. Dumais et al. (2001) developed and tested interfaces for presenting search results and found support for the suggestion that spatial grouping or the map display is an important feature used by skilled searchers. This interface organized results into spatial categories. The interface was similar to a map display. Under each category, relevant web pages to the query were listed. Additional pages are displayed by expanding the category. Dumais et al. (2001) tested the category interface through different experiments by adding or removing contextual information and adding or removing page titles. In all cases, the category interfaces were faster than the list interface in answering queries.

Now that we know the process of visualization, visualization formats should be designed to utilize features or techniques that can be automatically processed as well as provide support for search (Card et al., 1999). The next section reviews the usefulness of cues in visualization.

3.2 Visualization Techniques

Much of the information users receive is symbolic, consisting of numbers and texts. Processing symbolic information requires effort because it involves rule-based reasoning, where data is abstracted into values that have meaning for individuals (Sloman, 1996). However, individuals have great visual and spatial skills, allowing them to detect visual characteristics. According to Shneiderman (1996), “Humans have remarkable perceptual abilities that are greatly under-utilized in current designs. Users can scan, recognize, and recall images rapidly, and can detect changes in size, color, shape, movement, or texture” (p. 337). This enables individuals to retrieve information using visual cues (Kosslyn, 1994). These cues can be used to enable users to process information more effectively.

Card et al. (1999) lists several visual features or cues that can be automatically processed by humans. These are numbers, length, size, color, intensity, flicker, curvature, and direction of motion. Automatic processing is described as "superficial, parallel, can be processed nonfoveally has high capacity, is fast, cannot be inhibited, is independent of load, unconscious, and characterized by targets 'popping out' during search" (p. 25). To process "nonfoveally" is to do so visually but without eye movement. As our eyes move from one point to the next, nonfoveal vision sends back a preview of the next image (Coe, 1996). Animation is a feature that can be processed easily by nonfoveal vision. We plan to use animation as the primary cue in this study in order to emphasize relevant information within the display.

There are other various visualization techniques such as color, length, zoom and motion that can be incorporated into an information format to help individuals find results quickly by reducing information overload. In the next sections, we will review the different visualization techniques or cues as they apply to this study.

3.2.1 Color

Hoadley (1990) states "color is a subtle variable that can significantly enhance a decision maker's ability to extract information" (p. 125). Benbasat and Dexter (1986) found that color helped decision makers complete the decision-making task when under a more rigorous time constraint. Montazemi and Wang (1989) found that color improves decision quality especially for information users with a field-dependent personality. They also found that multi-color features were more useful than mono-color features. Several search engine interfaces have used color to present different categories to enhance users' comprehension. For example, Grokker.com provides a visualization of search results through a map display, which uses color to differentiate between groups and subgroups of results.

3.2.2 Size

Size is another useful cue utilized in various visual formats. Crapo et al. (2000) state size, motion, color, intensity, intersection, closure, orientation, and distance can be processed without conscious effort. These features seem to "pop out" in the visual interface (Healey et al., 1996). Percy and Rossiter (1983) found individuals' attitudes towards purchasing a product increased when viewing large pictures of the product versus smaller pictures of the same product. Font size has also been studied to identify how individuals read faster with larger font size than smaller font size (Chan and Lee, 2005). Grokker.com also uses size in its display to indicate the number of documents within each category.

3.2.3 Animation

Animation is defined as autonomous motions of representations (Nakakoji et al., 2001) or a series of rapidly changing computer screen displays that represent the illusion of movement (Phillips and Lee, 2005; ChanLin, 2000). Using animation as a cue can help individuals filter through the information more quickly by drawing users' attention. For example, animation is among the most prominent attention-getting devices used in web advertising (Sundar and Kalyanaraman, 2004).

Conventional animation tends to show information in a continuous format so that the entire presentation is displayed without any breaks (Mayer and Chandler, 2001). Motion is the key component of animation (Rieber, 1991) and in the visual area attracts attention (Hong et al., 2004; Lang et al., 2002). As our attention is drawn to certain stimuli, animation influences how well we perceive, recall, and act on information. Objects or information that does not receive attention fall outside our understanding and therefore have little influence on performance (Proctor and van Zandt, 1994; Hong et al., 2004).

In the early 1960s, simple animation was used in system interface design by providing blink coding. Blink coding is used to indicate an urgent need for user attention or to indicate the active location for data entry. Using simple animation in the interface provides attention to a particular part of the screen due to its visual distinctiveness (Hong et al., 2004).

Previous research in using animation on the Web has investigated flashing words (Heo et al., 2001; Li and Burkovac, 1999), speed of animation (Sundar and Kalyanaraman, 2004), interactive animated characters (Phillips and Lee, 2005), and effects of animation on web performance (Zhang, 2000). Zhang (2000) studied the effects of animation on information search performance. This study focused on how animation could be distraction when the user was searching for a string of letters. As the user was searching for letters, irrelevant animation on the screen was displayed. She found that irrelevant animation could distract the user's attention, therefore reducing the performance of information seeking. Craig et al. (2002) found that animation conditions improved performance by directing the learner's attention to specific elements of the visual display.

Nakakoji et al. (2001) studied the effects of animated visualization in exploratory data analysis tasks. They studied the effects of animations among tables, graphs, and animated graphs. They found that animations work effectively to view data in different viewpoints as well as focus on transitions of values of time-based data. In a second experiment, Nakakoji et al. (2001) also developed an interactive animated visualization environment that visualized the evolution of a programming library. They found that users were able to identify data points where values change as well as understand the data intuitively (Nakakoji et al., 2001).

Animation can be an important feature in interface design, but empirical research in this area is limited in the IS domain (Hong et al., 2004). According to a review by Hong et al. (2004) of human computer interaction (HCI) literature, animation is often adopted in IS for three functions: 1) "look and feel," e.g., entertainment (Thomas and Calder, 2001); 2) information visualization to increase comprehension (Mackinlay et al., 1994); and 3) attracting users' attention to specific information on the screen (Nielsen, 2000). Nakakoji et al. (2001) state that animated information systems visualize abstract data and represent changes of values in data over time using motions. The goal of these systems is to help users in analyzing data and making decisions by uncovering “hidden” meanings (Nakakoji et al., 2001).

Search engines have arbitrarily used animation in visualizing search results; however, this may not be useful for individual comprehension. For example, KartOO uses animation in their visualization to show that the search engine is “thinking” while search results are processed. In addition, as the user places her mouse over a particular document, the document is highlighted. However, because KartOO’s use of animation has no meaning to the user, it can be confusing.

3.3 Summary

In this chapter we discussed the process of visualization and the different cues that may be useful in a visualization format. We reviewed how animation as a cue is rarely utilized in current visualizations; however, animation can be effective in drawing attention to information. In the next chapter, we draw on information foraging theory to understand how individuals seek information using cues. We also discuss motion effect theories and the construct of distinctiveness to further understand animation as a useful cue. We also review cognitive fit theory as a framework to evaluate which different formats match tasks completed on search engines on the Internet for better performance.

CHAPTER 4

THEORETICAL FRAMEWORK

4.1 How do Individuals look for Information?

4.1.1 Information Foraging Theory

Information foraging (IF) theory was developed to explain human information seeking and sense-making behavior (Chi et al., 2000; Chi et al., 2001). IF theory deals with understanding how an individual uses strategies and technologies to seek, gather, and use information when there is a vast amount of information in the environment (Pirolli, 2003; Card et al., 2001). Its main focus is finding more information while expending less time and energy. IF theory describes the process individuals use to evaluate when to stop searching the available resources for relevant information.

Three core concepts in IF theory are information patches, information scent, and information diet. Information patches are similar to an individual's information needs that reside in piles of documents, results, file drawers, or various on-line resources (Card et al., 2001). Users navigate through these information patches to find relevant information. Often users have to navigate through more than one patch (i.e., from one web site to another or from one search engine to another) to find useful information. When beginning to navigate through a patch, the information is plentiful. However, with continued searching, the quantity and quality of the information begins to diminish (Gattis, 2002).

At this time, users are faced with a decision whether to keep searching or to stop browsing through the patch and move to another patch.

To help with this decision, individuals depend upon information scent. Information scent characterizes the individual's use of environmental cues in judging which information sources are important while navigating through an information space (Pirolli, 2003). It is the "imperfect perception of the value, cost, or access path of information sources obtained from proximal cues, such as www links" (Card et al., 2001, p. 499). For example, on a web page, information scent is delivered by a descriptor of the page, images, headings, or other cues. As discussed in the previous section, the cues of color and size are important scents that have been utilized in current visualizations to help individuals make decisions on whether information is relevant. In addition to color and size, there are other scents that can be useful to help make decisions.

These decisions help to ensure that individuals are maximizing their information diet. Information diet refers to the specific types of data or resources they select from all the possible data sources (Gattis, 2002). In most cases, information varies in quality and availability, and individuals, with their limited time, have to make the right decision whether to stay with one patch of information or move on. By choosing one source, they forgo other information sources in other patches, potentially resulting in a missed opportunity. Therefore, while navigating information patches, it is important for individuals to make good decisions in the data sources they choose for their diet.

Gattis (2002) used IF theory along with strategic planning theory to describe individuals' behavior in finding information. She found that IF theory helped explain how users identify search goals and allocate their time and energy for maximum gain.

She then used strategic planning theory to describe how humans achieve social goals by interacting with co-workers to obtain additional data. Card et al. (2001) proposed a protocol analysis methodology of user behavior on the Web using information scent as a driver. Participants were provided with six different browsing tasks to perform on the Web. They found that information scent was an important factor in how individuals found their results. Card et al. (2001) found that individuals initially found information scent to be high, but when scent became low, they switched to another page or search engine. "The idea is that a user is assessing the potential rewards of foraging at a site based on information scent" (Card et al., 2001). As long as the potential reward or finding the information is above some threshold, users stay at that site. Once the potential reward falls below the threshold, users move on.

Pirolli et al. (2001) used information foraging theory to compare the hyperbolic tree browser to a conventional browser. They found that while individuals were browsing the hyperbolic tree, they were affected by information scent and visual density. Pirolli et al. (2001) integrated visual attention with IF theory to find that individuals used scent to determine which objects can be searched with a hierarchical or a serial search based on the density of the groupings in the browser. Chi et al. (2000) developed a scent flow model to predict and analyze web site usability. Chi et al. (2001) developed algorithms to understand the concept of information scent by simulating a user's path through the Web. Through this algorithm, they were able to identify the kind of information the user was seeking. This information would help designers personalize web environments, create web sites based on information goals of users, as well as identify poor web site designs.

In practical terms, this theory can help explain which cues and design alternatives would be best suited for web pages or search engines. Rather than randomly create search engines with visualization techniques, IF theory can point designers to specific cues that will help individuals improve their performance, and determine which cues should be ignored (Pirolli, 2003). In an environment of information overload, the design problem is not how to collect more information but how to increase the amount of relevant information found. Since individuals have limited time and attention, they prefer to select designs that improve information gathering (Pirolli, 2003).

As discussed earlier, individuals typically forage for information on the Web by navigating through web pages via hyperlinks. These hyperlinks are presented to individuals through "some snippets of text or graphics called browsing cues" (Olston and Chi, 2003, p. 180). According to Olston and Chi (2003), individuals use browsing cues to access the "distal" content (the page at the other end of the link). Through information scent, individuals use perceptual cues to judge information sources and navigate to them while exploring and searching for information (Pirolli, 2003).

Depending on the strength of the scent, individuals are able to decide whether to exert time and energy navigating through the information or to make better use of resources elsewhere. Browsing cues, such as hyperlinks, have limited scent since they lack enough information to provide individuals with sufficient guidance to forage through results. According to Olston and Chi (2003), these limitations are caused by three things. First, inappropriate cues can lead to poor linking of results. Second, cues do not provide enough information of the web page. Lastly, browsing cues are not customized to the individual's information goals.

Color, size and animation are useful scents or cues. Animation has the potential to have a strong information scent, thereby serving as an effective visual cue enabling users to find relevant information in a short time. IF theory helps us understand how enhancing visualizations of search engine results using appropriate cues, such as animation, can help individuals more easily find relevant information. Next, we will discuss the cues of color, size and animation as useful cues or scents. Since the focus of this study is on animation, we will discuss motion effect theories to explain how animation can draw attention to relevant information, therefore serving as a useful and strong visual cue for search query results.

4.2 How can Animation be a Useful Cue?

4.2.1 Motion Effect Theories and the construct of Distinctiveness

Motion effect theories assume that individuals have an inherent preference for moving objects. When users are "exposed to moving images, they focus their attention on the source of motion and process relevant information" (Sundar and Kalyanaraman, 2004, p. 8). According to these theories, when users are exposed to a visually surprising object, they focus on the source of the animation and stop all other unnecessary activities (Heo and Sundar, 2000).

Based on the prior discussion, presenting results from search engines in a visual format may reduce information overload by providing visual cues to locate relevant documents. Information foraging theory suggests that cues such as animation may help individuals find relevant documents more quickly. Motion effect theories provide the rationale that animation is a useful cue because it draws an individual's attention. These theories help us understand how using animation in a search engine interface can help individuals more effectively find relevant information. However, the ease in which individuals find information may also depend on matching particular visualization formats to particular information search tasks. In the following paragraph, we present the construct of distinctiveness to support the notion that individuals will draw more attention to certain features, such as animation, if it is different from the rest of the display.

A stimulus can be considered “distinctive” if it has unique features that distinguish it from the rest of the stimuli in an individual’s visual field (Phillips and Lee, 2005). The construct of distinctiveness establishes that the more distinctive something is, the more likely it is to be recognized and remembered. Research in the construct of distinctiveness has focused on animated banner ads in Internet advertising (Kim et al., 2003; Yoo et al., 2003; Li and Burkovac, 1999). Li and Bukovac (1999) believed that larger banner ads and banner ads with animation would result in shorter response times and higher recall, than smaller ads and those without animation. The study showed that animated banner ads were clicked on much more quickly often, and did result in higher recall than static banner ads.

There is support for animation as a “distinctive” feature on web pages. Animated banner ads draw greater attention than do static banner ads (Kim et al., 2003; Yoo et al., 2003). Recall and recognition of the subject is also higher for animated ads compared to static ads (Li and Burkovac, 1999; Yoo et al., 2003). Animated banner ads are distinctive from the text that surrounds them; therefore, the animated ads draw more attention from the individuals, resulting better recall.

In contrast, there is research that shows animated ads do not have a significant effect on consumers compared to static ads. Heo et al. (2001), using eye-tracking movements to measure attention, found there were no benefits of attention for animated ads over static ads. Phillips and Lee (2005) found that animated web ads tend to distract users from their original goal if the users' intent was not to make a purchase. There are also limited studies in other areas of research on the use of animation. Turetken and Sharda (2007) state that few web visualizations employ animation despite its potential usefulness. Based on the construct of distinctiveness, individuals should respond positively to animation as a cue different from the rest of the visual format, as it would draw users' attention and direct them to relevant search engine results.

Based on the prior discussion, presenting search engine results in a visual format may reduce information overload by providing visual cues to locate relevant documents. Information foraging theory suggests that cues such as animation may help individuals find relevant documents more quickly. The construct of distinctiveness provides the rationale that animation is a useful cue because it draws an individual's attention. These theories help us understand how using animation in a search engine interface can help individuals more effectively find relevant information. However, the ease in which individuals find information may also depend on matching particular visualization formats to particular information search tasks. In the following section, we discuss cognitive fit theory, which explains how matching format to tasks lead to better performance.

4.3 Framework to Evaluate Formats and Tasks

4.3.1 Cognitive Fit Theory

Cognitive Fit Theory (CFT) was developed to explain how the appropriateness of presentation format to decision-making tasks could affect individual's problem-solving performance (Vessey and Galletta, 1991; Vessey, 1991). CFT suggests that when there is a mismatch between the information format and the task, the individual will invest more cognitive effort in the decision-making process because they need to adjust their mental representation to accommodate the mismatch. Cognitive effort refers to the "psychological cost of performing the task of obtaining and processing the relevant information in order to arrive at one's decision" (Hong et al., 2004, p. 159; Pereira, 2000). If individuals need to adjust their mental representation to make decisions, this will be result in lower performance. Hong et al. (2004) suggest users will also have a better attitude toward the web site when there is a match between information format and task.

There are two types of tasks described in CFT: spatial and symbolic. Spatial tasks refer to tasks that require individuals to make associations about relationships between the data. Spatial tasks are similar to open-ended tasks, where results need to be obtained from various sources and individuals need to understand the material to provide a valid answer. Symbolic tasks are similar to closed-ended tasks and involve extracting discrete data values.

According to CFT, spatial tasks are best supported by spatial formats and symbolic tasks are best supported by symbolic formats. Symbolic formats are those that present data values that can be calculated and computed easily, similar to a standard search results page. Spatial formats are those that present relationships in the data rather than actual values, similar to a visualization format.

Based on CFT, a better fit between format and task should result in improved performance, which can be measured in various ways (Hong et al., 2004; Agarwal et al., 1991; Vessey and Galletta, 1991). Hong et al. (2004) measured performance by the web site that subjects chose while shopping online and their attitude towards the web site. Agarwal et al. (1996) measured success by which methodology was chosen by the participant to perform a system analysis and design task. The study involved object-oriented and process-oriented methodologies of business information-processing problems. Subjects used either object-oriented tools or process-oriented tools to complete the tasks. They found superior performance for process-oriented (PO) tasks when using the process-oriented (PO) tool, but not for object-oriented (OO) tasks when using the object-oriented (OO) tool. Despite this, users were able to find more minor relationships using the OO tool over the PO tool. This could be attributed to the fact that OO tools support a graphical representation that helped with problem solving for OO tasks. They conclude that the lack of improved performance for OO tasks using OO tools is that humans have an innate tendency to use procedural or process-oriented approaches to problem solving.

Vessey and Galletta (1991) examined CFT within the domain of graphs versus tables. The tasks involved problems of a bookkeeper who had a number of bank accounts under his control. The information formats used in their study were line graphs and two-dimensional tables, and used two performance measures: time taken and accuracy of completed tasks. Spatial tasks included finding which month the difference between deposits and withdrawals were greatest and symbolic tasks included providing the amounts of withdrawals and deposits in various months. They found for spatial tasks, users solved problems more quickly with graphs than users with tables. However, users solving problems with tables were more accurate. They also found that for symbolic tasks, tables resulted in faster and more accurate performance than graphs. In addition, they found that matching problem-solving skills to the representation and task leads to further improved performance.

Many researchers have used CFT to demonstrate the differences between problem-solving tasks on graphs and tables (Vessey and Galletta, 1991; Vessey, 1991). Vessey (1991) suggest that graphs fit more with tasks that emphasize spatial process because graphs present relationships in data rather than values. Tables fit more with tasks that emphasize symbolic process because tables present discrete data values that can be used easily for computations.

However, Vessey (1991) states, “cognitive fit is not restricted to the graphs versus table domain. It can be applied to any domain where there is sufficient information to permit analysis of the tasks to be performed” (p. 234). Agarwal et al. (1996) used CFT in to study object and process methodologies as applied to object-oriented and process-oriented tasks. Dennis and Carte (1998) used CFT to understand decision-making processes on a spatial decision support system (SDSS), a type of geographic information system. Table 1 summarizes the different domains to which CFT has been applied.

Dennis and Carte (1998) state the success of decisions depends on not only the fit between the presentation and task but also the processes performed by the individual. Decision processes are affected by perceived cost and benefits to users. Users want to choose processes that have lowest cost and highest benefits (Russo and Dosher, 1983). Dennis and Carte (1998) state that decision makers choose processes that are most appropriate to the format of the data they are given. When using map-based presentations, decision makers use less cognitive effort since they do not need to convert spatial data into precise numeric data to make a decision.

Several studies have examined other factors that may influence cognitive effort's effect on performance. In addition to format and task, Vessey and Galletta (1991) also used "skill" as another construct in their study. They found that seeking information-processing skills that support a particular task had a greater effect on performance when format and task matched. Sinha and Vessey (1992) tested an extended model of cognitive fit that involved a "problem-solving tool" as an additional factor to determine performance. They found the effects of a match between the task and the problem-solving tool (programming language) outweighed the match between representation and task.

4.4 Summary

In this chapter, we reviewed information foraging theory to explain how individuals look for relevant information when there are vast amounts of information. Individuals use a scent or a cue to help make decisions on which set of information are relevant. We reviewed motion effect theories and the construct of distinctiveness to show that animation is a useful scent or cue on a visualization format to help individuals find relevant results.

Table 1. Cognitive Fit Theory in Different Domains

Authors	Tasks	Format	Performance Measures	Domain
Vessey and Galletta (1991)	Spatial Symbolic	Spatial (Graphs) Symbolic (Tables)	Accuracy Time/efficiency	Graphs versus Tables
Mahoney et al. (2001)	Spatial Symbolic	Spatial (Graphs) Symbolic (Tables)	Accuracy Time/efficiency	Graphs versus Tables
Chan (2001)	Spatial	Spatial (Graphs) Symbolic (Tables)	Accuracy	Graphs versus Tables
Hong et al. (2004-5)	Searching Browsing	List Matrix	Time Recall Attitude Cognitive decision effort Cognitive convenience	Internet shopping
Huang et al. (2006)	Association Compare Distinguish Rank Cluster	SOM MDS	Efficiency Effectiveness	Visualization of field experts
Dennis and Carte (1998)	Geographic containment (symbolic) Geographic adjacent (spatial)	File Reader (textual) Atlas-Graphics (GIS)	Decision process Decision accuracy Decision Time	GIS
Speier, Vessey, and Valacich (2003)	Symbolic	Spatial (Graphs) Symbolic (Tables)	Accuracy Time	Interruptions in work tasks
Umanath and Vessey (1994)	Holistic	Holistic Spatial Symbolic	Accuracy Time	Accounting

CHAPTER 5

DEVELOPMENT OF PROTOTYPES

This section discusses the development of the four prototypes utilized in this study. Prototypes 1 and 2 are created with an HTML interface and prototypes 3 and 4 are created with a Macromedia Flash interface. Below we discuss the three components required for development of the prototypes: retrieve results, organize results, and present results. A high-level activity chart of the process of developing the prototypes is provided in Figure 7.

5.1 Part A: Retrieve Results

In Part A, we obtained the search results used for this experiment. The search results should be the same for each query across the four prototypes used in the study. This experimental control ensures that performance does not depend on the search results of a particular query. We controlled for all factors except for information format, which is different for each prototype. Therefore, we can ensure that performance depends only on the varying factor, information format.

To obtain the search results, we develop a server-based ASP.NET application using C#. All results are obtained from the MSN Live Search engine. We used MSN's Windows Live search engine because its straightforward API allowed us to easily construct our prototypes and more easily retrieve the search engine results (Google's API was more restrictive; for example, it would only allow us to retrieve 10 results at a time). The C# program used MSN Live's Web Services API to obtain approximately 300 results for each query. The results are parsed into "title", "URL", and "snippet" (a snippet is a fragment of text from the body of the document). In addition, we determined each search result's ranking or relevancy by the order of the results obtained from MSN. These parsed results are saved as an XML file.

Table 2. Summary of Steps in "Part A" of the Prototype Systems

Part A: Retrieve Results
C# program uses Windows Live Search Web Services
Obtain 300 results from MSN
Parse web results of title, URL, and snippet
Parsed results saved as an XML file on the server

5.2 Part B: Organize Results

In Part B, we utilized Carrot^2, an open source clustering engine. Carrot^2 has obtains the search results and automatically organizes them in clusters. The architecture of Carrot^2 is based on three components: input, filter and visualization. The input component provides search results for clustering. The filter component transforms the results through algorithms, and the visualization component provides the clustered results for the user. For this study, we do not utilize the visualization of the clustered results, since we created our own visualization formats.

Since the Carrot^{2}API is written in Java, we used a Java application to pass the XML files containing the saved MSN Live Search results to Carrot^2. Carrot^2 provides an option of five different clustering algorithms. For this study, we used the Lingo clustering algorithm from Carrot^2 since it provided the most understandable labels for the clusters.

Carrot^2 provides a file containing the cluster structure, the name of the cluster and the contents of each cluster. Another Java application reads the data structure file from Carrot^2 and saves an output file in XML with the cluster information (including its relevancy). Therefore, the search results as XML files from Part A as well as the XML clustering output file from Part B are saved on the server to be accessed later for creating the display.

We determined each search result's ranking or relevancy to the query by the order of the results in the search list obtained from MSN. Relevancy is a measure of how closely an individual search result matches the original query. The relevancy is calculated using the mean reciprocal ranking (MRR). The calculation for MRR is as follows:

$$MRR = \frac{\sum_{i=1}^{n}\left(\frac{1}{Rank_i}\right)}{n}$$

where

$Rank_i$ is the rank of the i^{th} search result in the cluster and n is the number of results within the cluster.

Table 3. Summary of Steps in “Part B” of the Prototype Systems

Part B: Organize Results
Java program feeds saved XML files into Carrot^2
Carrot^2 clusters results
Java program reads the cluster structure, name of cluster, and results
Java program calculates the relevancy for each cluster
Java creates an XML file of cluster information and the XML file is saved to server

5.3 Part C: Present Results

In Part C, we discuss the presentation formats of the results obtained from Part A and Part B for each prototype in the study. For prototype 1, a C# application displays search results in a ranked list as an HTML page (See Figure 3). When the user clicks on the link, the target web page displays in a separate window.

For prototype 2, textual with categories, a C# application parses the XML file obtained from Part B and displays the cluster structure in a tree, with the first-level branches of that tree containing the name of the clusters (See Figure 4). When a user clicks on a cluster name, the interface will display results within that cluster. The search results will be displayed by its URL and title. At this level when a user clicks on an item, another window will appear with the web page.

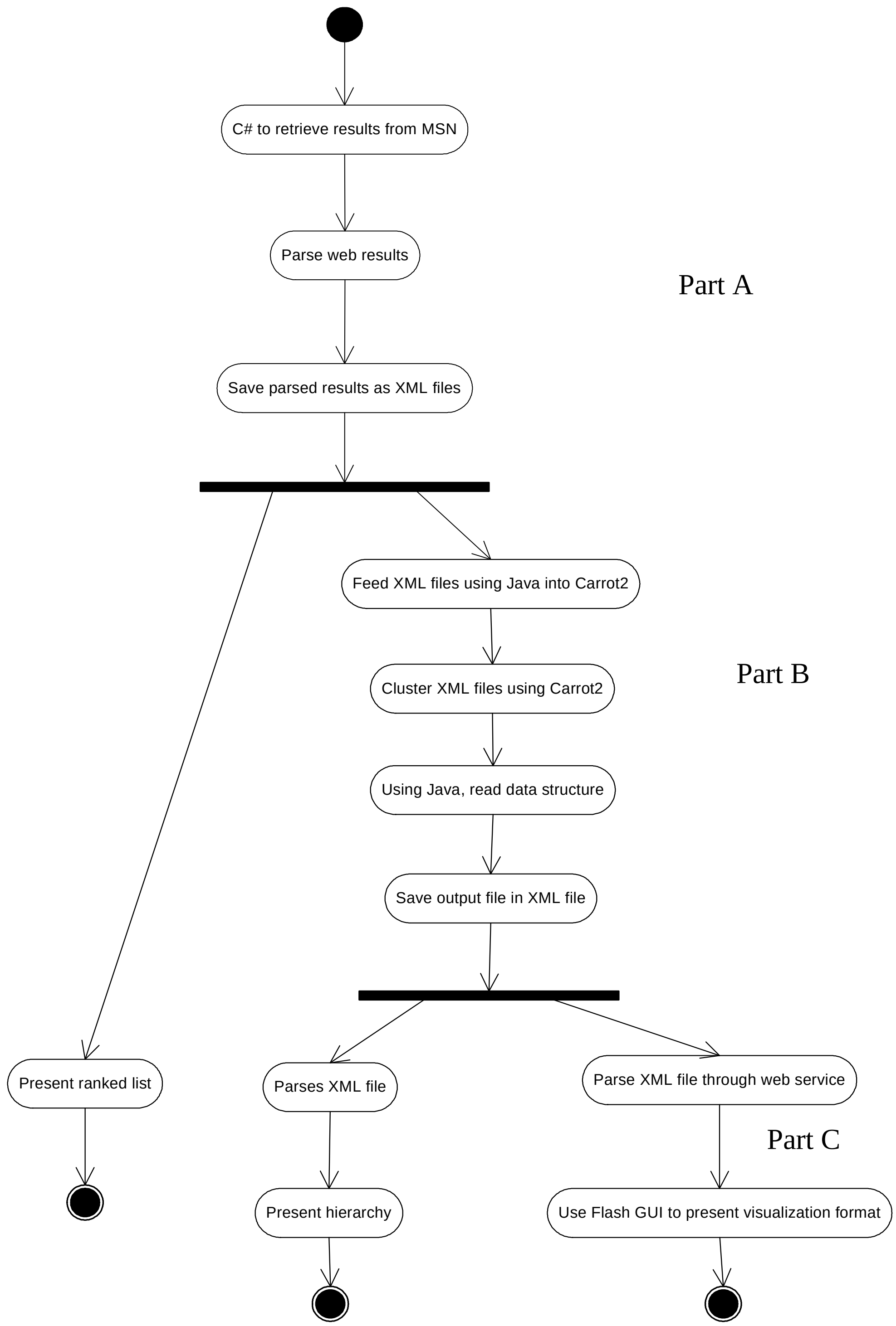


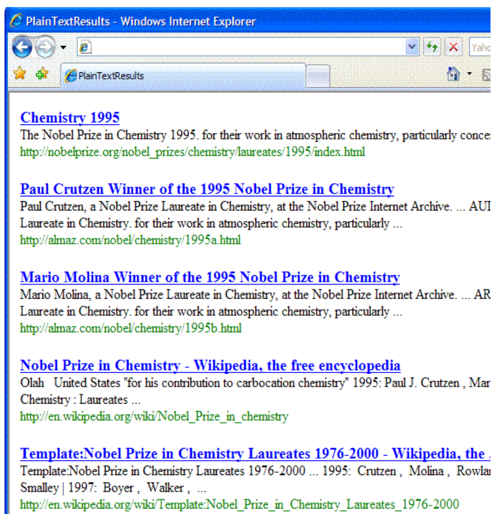
Figure 2. Activity Chart of Design/Development Process

For prototype 3 (non-animated visualization – See Figure 5) and prototype 4 (animated visualization – See Figure 6), we used a Macromedia Flash-based application to visualize the results. In order to display the clustered tree in the Flash application, a web service (also written in C#) parsed the XML file obtained in Part B into a series of strings (one for each cluster) that contain XML-formatted data tree. The Flash program parsed each string and feeds that data tree structure into a Macromedia Flash-based application tree control to display the hierarchical structure required for each circle or cluster. The Flash program displays each cluster as a circle in a map format. When a user clicks on a circle, the interface displays another window with all the results within that cluster, which is the hierarchical structure (reflected in the XML document) displayed through the Macromedia Flash tree control. The results are displayed by its title. When a user selects a particular title, the program will look up the URL from that cluster's XML file, and display the web page in a new window. For prototypes 3 and 4, the user is able to view the map display of the results as well as the web page in separate windows.

Prototypes 2, 3, and 4 also provide the user with the relevancy of each cluster of results to the query. For prototype 2, textual with categories, we display the relevancy score next to the cluster name (See Figure 4). For prototype 3, non-animated visualization, we display the relevance for each category within its circle (See Figure 5). For prototype 4, animated visualization, we communicate the relevance for each circle through blinking (See Figure 6). Circles blink based on its relevancy to the query. The most relevant cluster or circle blink the fastest, with the second most relevant cluster blinking slower and third most relevant cluster blinking even more slowly.

Table 4. Summary of Steps in "Part C" of the Prototype Systems

Part C: Present Results
C# program presents ranked results from Part A (Prototype 1)
C# program parses XML file from Part B and presents hierarchy of tree (Prototype 2)
C# web service parses XML file into a series of strings
Flash code parses strings and feeds into Flash tree control
Flash program to visualize results into map (Prototype 3 and Prototype 4)
Based on the prototype, relevancy will either be displayed through text or blinking.



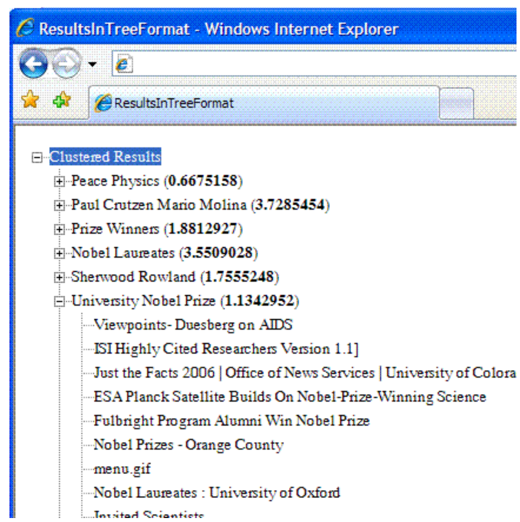


Figure 3. First Format: Standard Textual

Figure 4. Second Format: Clustered Textual

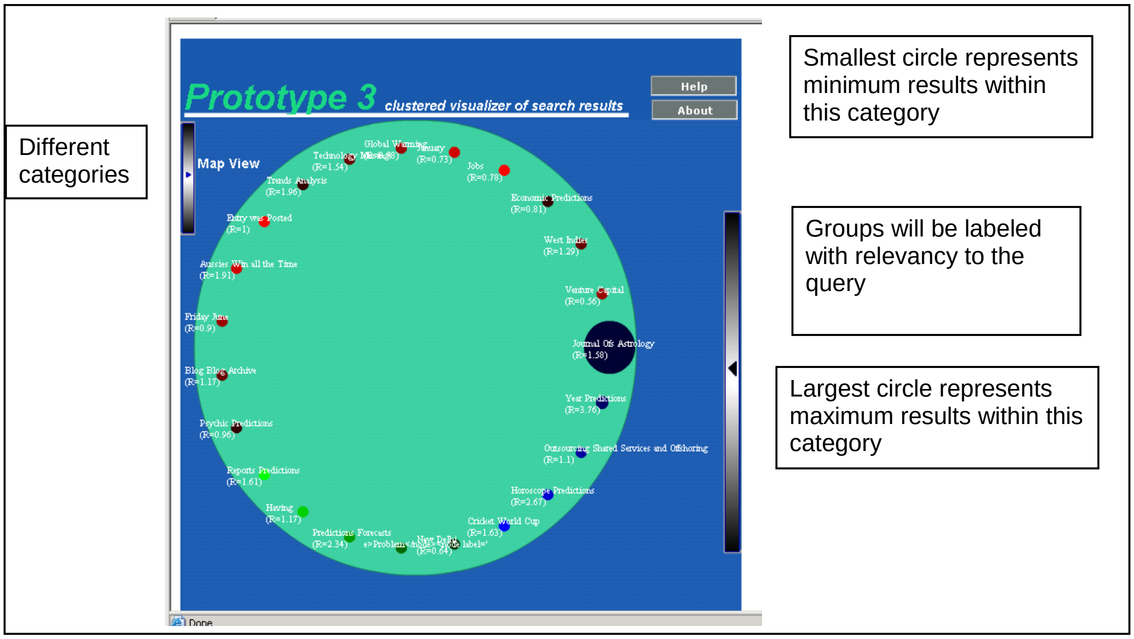


Figure 5. Third Format: Non-animated visualization

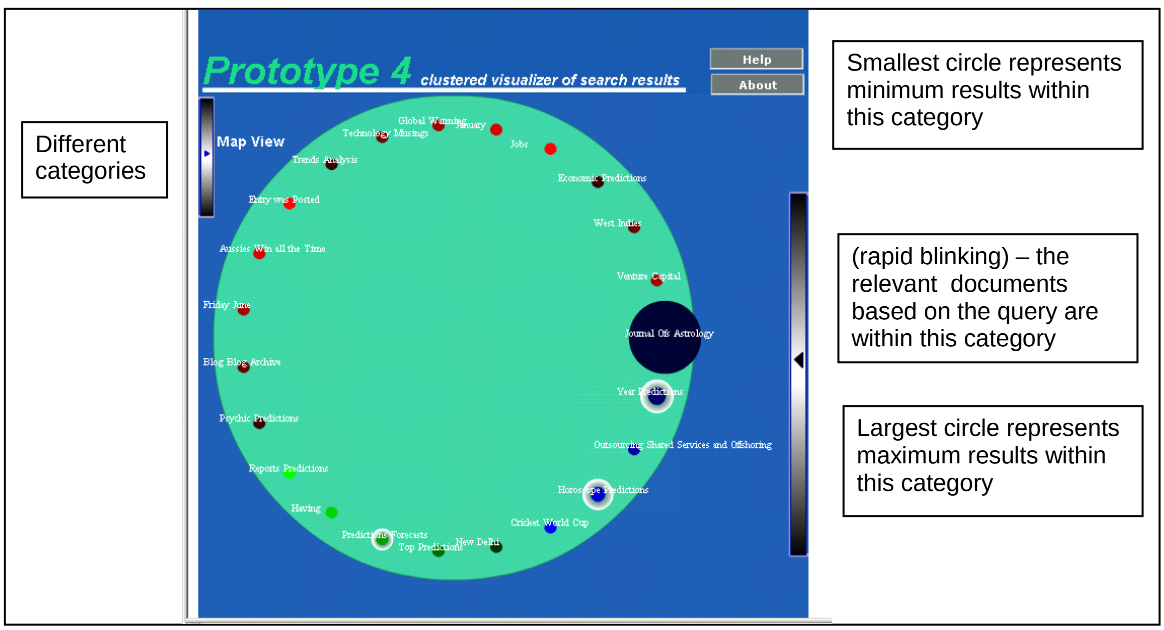


Figure 6. Fourth Format: Animated visualization

5.4 Summary

In this chapter, we discussed the steps that were taken to create the different prototypes (formats) that are evaluated in this study. In the following chapter, we review the proposed hypotheses.

CHAPTER 6

CONCEPTUALIZATION OF FIT

6.1 Tasks: Information-seeking Tasks

According to information retrieval literature, searching and browsing are two general methods performed to seek information on the Internet (McDonald and Chen, 2006). Marchionini and Shniederman (1988) define searching as more focused, while browsing is described as "an exploratory, information-seeking strategy that depends on serendipity" (p. 71). Kuhlthau (1991) found that individuals engage in general browsing of the topic until understanding of the topic occurs, after which a more directed search takes place. Cove and Walsh (1998) used a three-stage model with 'knowledge of the goal' as the primary factor in the information-seeking stage. However, when the information needed is less clear, a general and "serendipitous" browsing pattern is performed. Individuals look for general topics in the beginning of the information-seeking process, and towards the end of the search, they look for more specific information regarding the topic.

In this study, we investigate the effect of visual cues (specifically, animation) on the presentation of search engine results through both open-ended and closed-ended tasks. Our definition of closed-ended tasks is tasks with specific objectives and has a well-defined answer. We define open-ended tasks as tasks that involve gathering information on a topic in order to answer a question where the answer is not well-defined.

6.2 Information Format and Visualization Techniques

"Information format" is defined as the presentation and organization of information (Hong et al., 2004; Cooper-Martin, 1993). Different information formats, such as tables and graphics, emphasize different types of information and problem-solving processes (Hong et al., 2004). This study investigates two broad types of information formats: textual and visual. Textual refers to a character-based representation of words and numbers, similar to what one would find on a Google search result page. The arrangement of results is from top to bottom and can be read from left to right down the page. Textual representations of search engine results are symbolic (as discussed in section 3.3) because they simply present the results without providing information regarding the relationships between them.

Visualization refers to a visual representation of words and numbers through visual cues such as symbols and colors. The arrangement of results varies based on how the cues are displayed on the format. These representations of search engine results are spatial because they show relationships among the results as well as the results itself. It shows this additional information through a graphical representation.

This study uses four specific information formats: standard textual, clustered textual, non-animated visualization, and animated visualization. We discussed the development of the four prototypes in the previous chapter.

6.3 Hypothesis Development

6.3.1 Cognitive Effort

Presenting results in a particular visualization format may reduce information overload by providing cues to locate relevant documents for particular tasks. Information foraging theory suggests that scent is provided through cues that may further help individuals find relevant documents more quickly. Based on CFT, the ease with which individuals find information may also depend on matching particular formats to particular information search tasks to reduce cognitive effort. In this section, we discuss the rationale for the hypotheses that describe the fit between format and tasks, beginning with closed-ended tasks followed by open-ended tasks.

Based on CFT, there will be higher performance if the task fits the format (Vessey and Galletta, 1991). Previous research applied CFT to symbolic and spatial tasks stating that symbolic tasks fit with symbolic formats and spatial tasks fit with spatial formats to reduce cognitive effort (Refer to Table 1). Closed-ended tasks are symbolic in nature. They involve extracting discrete data values or specific results. When individuals are performing closed-ended tasks, the answer is narrowly defined and their search has a specific goal, reflected in a specific query. For symbolic tasks or closed-ended tasks, there is a precise answer that can be found more easily using a symbolic format. Textual formats (the “standard” format mentioned in the previous section) are similar to symbolic formats, which emphasize discrete data values or results. They do not represent relationships within the data. Since standard formats are used to extract specific data or results, these formats would best fit with closed-ended tasks, where the goal is to find a specific result. Using a standard textual presentation, the most relevant results (those that match the text of the query) are found easily by looking to the top of the ordered list. For closed-ended tasks completed on standard format, the available cues such as titles of websites and rank in the list would further help users find the results quicker.

A closed-ended task, such as “Who won the 1979 Nobel Peace Prize?” that is completed using a search engine that presents a text-based list of results (like Google) would present various results in a ranked ordered list, similar to the standard format. There is only one correct answer to complete the task. Based on information foraging theory, there is a scent that is provided from the cues such as titles of the websites and rank that would help users pick a relevant result, which is usually at or near the top of the ranked list. To confirm the answer, he might also select the second website in the ranked list. The entire task would be completed within a few minutes without having to browse through irrelevant websites. Because there would be a fit between the format and task, subjects should require less cognitive effort when using a standard format over one where the results are grouped into categories. This is because subjects using a standard format do not have to understand the structure of and navigate through the grouped categorized set of results to find answers, which would result in a higher level cognitive effort required in order to complete the task.

If the closed-ended task stated above, “Who won the 1979 Nobel Peace Prize?” was completed using a search engine that presents its results in a visual map (like Grokker), the user would have to navigate the display to find results. Based on information foraging theory, individuals would utilize scents provided from various cues such as titles, color, size, relevancy, or animation to find a relevant category that might hold the answer (Pirolli, 2003). The user would then have to select websites within the category to find the correct answer. The selected category may also contain irrelevant results, where the subject may select while trying to find the correct result. The amount of time that it would take an individual to complete a closed-ended task on a visualization format is slightly higher than finding the answer in the first few results on a standard format. Similarly, subjects using more complex, sophisticated formats to navigate the results should require increasing levels of cognitive effort to further understand the results. Cognitive effort required gradually increases to complete closed-ended tasks for each format as depicted through the arrow in Table 5 that moves from standard format towards animated visualization format. Therefore, we hypothesize:

H1a: For closed-ended tasks, there will be less cognitive effort required when using a textual format over a visualization format.

H1b: For closed-ended tasks, there will be less cognitive effort required when using a standard textual format over a clustered textual format.

H1c: For closed-ended tasks, there will be less cognitive effort required when using a non-animated visualization format over an animated visualization format.

Based on CFT, spatial tasks fit with spatial formats to reduce cognitive effort (Vessey and Galletta, 1991; Vessey, 1991). Spatial tasks refer to tasks that require individuals to make associations about relationships between the data or results (Vessey and Galletta, 1991; Vessey, 1991). Spatial tasks are similar to open-ended tasks, where results need to be obtained from various sources and individuals need to understand the material to provide a valid answer. Open-ended tasks involve general topics, such as "How do I obtain a Nobel Peace Prize?" or "What is the best digital camera?" When individuals are performing open-ended tasks, the answer is not as well defined as in closed-ended tasks. Kuhlthau (1991) found that individuals engage in general browsing of the topic until understanding of the topic occurs, after which a more directed search takes place. Cove and Walsh (1998) used a three-stage model with ‘knowledge of the goal’ as the primary factor for broad tasks. When the information needed is less clear, a general and “serendipitous” browsing pattern is performed. Individuals look for general topics in the beginning of the information-seeking process, and towards the end of the search, they look for more specific information regarding the topic. These tasks involve viewing multiple results and gathering information from different sources to make a decision. There is no correct result to complete the task. The goal of answering open-ended task is finding results or making relationships from the results to gain understanding and knowledge to be able to make a good decision.

Spatial tasks can be easily completed on spatial formats, which present spatially related data points and emphasize relationships in the data (Vessey and Galletta, 1991; Vessey, 1991). In this study, data are the search engine results. Spatial formats are similar to the third and fourth “information visualization” formats described in the previous section. Since spatial formats or information visualization formats allow users to make relationships from the relevant results, there will less cognitive effort for individuals completing open-ended tasks on visualization formats instead of a standard format. Information visualization formats categorize similar results together so that individuals are able to view results that are related to each other. Individuals are able to see the whole group of results instead of just seeing parts of the results.

Based on information foraging theory, cues such as color, size, or animation on visualization formats can provide scents to individuals to help find results (Pirolli, 2003). Since animation is defined as a moving image, animation can be distinctive on a format where all other images are static (Phillips and Lee, 2005). Because of its distinctive characteristic, animation will bring users' attention to the categories that are animated and individuals are able to process that information quickly (Sundar and Kalyanaraman, 2004). In this study, the top relevant three categories in the information visualization format are animated. Based on information foraging theory, we believe that animation will provide a strong scent for users to find relevant categories quickly. By selecting the relevant category quickly, subjects would be able to browse similar results to gain knowledge on the open-ended task faster than having to browse through a standard format where results are dispersed throughout the list.

Animating the top three links is not the only format change. In an open-ended task, the user often needs to investigate other links. The links are usually related in some way. For instance, an open-ended task such as finding a camera would involve figuring out many categories of issues, such as budgets, types of photos to be taken, expectations, and experiences or skills. Grouping such categories and animating relevant categories, as in the animated visualization format, will facilitate discovering those issues more systematically and completely. The number of irrelevant results would be limited since similar results would be grouped together, further helping to browse through results to gain knowledge on the task. Therefore, they are able to synthesize multiple results in order to arrive at an answer.

If the above task was completed using a text-based format, the results would be presented in a ranked list, which are not ordered based on similar results. Because of this, individuals using a textual presentation may have to browse through many sites throughout the set of query results (and not necessarily those at the top of the list). It would be difficult to discover relationships from the results since relevant results to help gain knowledge on the topic would be dispersed throughout the list. This may result in a longer time to complete the task, and possibly frustration leading to higher cognitive effort and even quitting the task altogether (Roussinov and Chen, 2001).

Since open-ended tasks involve discovering relationships from the results, an information visualization format would fit the task more than a text-based format. Grouping related results into categories and a visual map should further help users navigate the results, and the use of animation as a cue to highlight relevance should help further still. Therefore, animated information visualization formats fit with open-ended tasks, resulting in a lower level of cognitive effort required. Cognitive effort required gradually increases to complete open-ended tasks for each format as depicted through the arrow in Table 5 that moves from animated visualization format towards standard format. Therefore, we hypothesize:

H1d: For open-ended tasks, there will be less cognitive effort required when using a visualization format over a textual format.

H1e: For open-ended tasks, there will be less cognitive effort required when using a clustered textual format over a standard textual format.

H1f: For open-ended tasks, there will be less cognitive effort required when using an animated visualization format over a non-animated visualization format.

We will measure cognitive effort required using Hong et al. (2004) instrument that combines cognitive decision effort and cognitive convenience (refer to Appendix C). Although this survey was developed for the e-commerce domain, we will rephrase the questions to suit the current context. When there is less cognitive effort required, there is a cognitive fit between the format and task as described in Table 5.

Table 5. The Effect of Matching Task and Format on Cognitive Fit

	Textual Format		Visualization Format	
	Standard textual	Textual with categories	Non-animated visualization	Animated visualization
Closed-ended task	Highest cognitive fit			Lowest cognitive fit
Open-ended task	Lowest cognitive fit			Highest Cognitive fit

The research model is represented in Figure 7. Based on the model, we propose when there is a match between the visualization format (textual versus animated) and information-seeking tasks (closed-ended versus open-ended), more positive outcomes will occur in performance and satisfaction. We will now discuss each of the dependent variables and the related hypothesis.

6.3.2 Performance

Based on CFT, individuals will achieve a higher level of performance if there is a match between information format and task. From our earlier discussion, we proposed that this type of match enables individuals to perform better by making better decisions regarding which information is relevant to their search. We will measure performance through effectiveness and efficiency (Chung et al., 2005; Turetken and Sharda, 2005; Roussinov and Chen, 2001; Vessey and Galletta, 1991).

6.3.2.1 Effectiveness

Effectiveness of the visual format is based on two factors: exactness and the F-value (Chung et al., 2005). Precision and recall are measured in the study and calculated to obtain a single measure for the F-value (Chung et al., 2005). Exactness refers to how well the visual format helps individuals find correct answers to closed-ended tasks. This will be measured by calculating how many correct answers are obtained, divided by the number of questions. Based on the Chung et al. (2005) study, exactness will be measured only on closed-ended tasks (Task 2) since they require specific answers.

According to Chung et al. (2005), precision and recall can only be measured on open-ended tasks (Task 1) since there are no specific answers, just those that are relevant to the decision process. Participants will be asked to record the titles and URLs of all websites related to the task. Precision measures how well the visual format helps individuals find pertinent results and avoid extraneous results in open-ended tasks. Precision will calculate the number of relevant results identified by the participant, divided by the number of all results obtained by the participant in open-ended tasks. Recall measures how well the visual format helps the individual find all the relevant results in open-ended tasks. In order to measure recall, we will evaluate the participant's responses against previously obtained responses. Prior to the start of the experiments, three experts that are not familiar with the study will perform each task using MSN and provide all possible answers to each task. Recall will then be calculated by the number of results obtained by the participant, divided by the number of results obtained by experts.

The formulas for exactness, precision, recall, and F-value are listed below (adapted from Chung et al., 2005):

$$\textit{Exactness} = \frac{\textit{Number of correctly answered questions}}{\textit{Total number of questions}}$$

$$\textit{Precision} = \frac{\textit{Number of relevant results identified by the participant}}{\textit{Number of all results identified by the participant}}$$

$$\textit{Recall} = \frac{\textit{Number of relevant results identified by the participant}}{\textit{Number of relevant results identified by experts}}$$

$$\textit{F-Value} = \frac{2 * \textit{Recall} * \textit{Precision}}{\textit{Recall} + \textit{Precision}}$$

When individuals are able to find the answers to their query without having to filter through unnecessary information, their performance improves. For open-ended tasks, the F-value will determine effectiveness; while for closed-ended tasks, exactness will determine effectiveness.

H2: Effectiveness will be higher when there is less cognitive effort required (i.e., when there is a match between the visualization format and the information-seeking task).

6.3.2.2 Efficiency

According to Chung et al. (2005), efficiency refers to the amount of time it took individuals to complete an information retrieval task. Individuals want to find relevant information quickly. They want to complete their task using search engines quickly without getting frustrated and quitting. Less time also indicates better visualization design of query results, which is desirable to web designers. Information search time was measured by recording the starting time and ending time it will take each individual to complete each task. Efficiency will be measured on all open-ended tasks separately from all closed-ended tasks performed on each format. Therefore, the less time it takes an individual to perform the task on a given format will increase efficiency.

H3: Subjects will take less time to complete tasks when there is less cognitive effort required.

6.3.3 Satisfaction

Satisfaction is defined as the amount of pleasure obtained from using a particular information format. If individuals are satisfied the visualization of query results that enabled them to complete their task, they will be more willing to use that tool again. DeLone and McLean (1992) state that one measure of IS success is user satisfaction. Turetken and Sharda (2005) use satisfaction as a measure to identify whether users were satisfied with a particular information format. Chung et al. (2005) measured usability, which was defined as how satisfied users were with the browsing method. In this study, we use satisfaction as a surrogate for usability to measure how users felt about the given information format. For measuring satisfaction of search engines, Turetken and Sharda (2005) adapted a multi-item scale from Stasko et al. (2000). This adapted satisfaction scale will be used in this study (refer to Appendix B). We believe when there is a match between information format and task; less cognitive effort will need to be invested, leading individuals to be more satisfied with the format.

H4: Satisfaction will be higher when there is less cognitive effort required.

Based on research in human resources management (Christen et al., 2006; Abdel-Halim, 1980), individuals that perform well at their job tend to be more satisfied with their job. If individuals feel that they are successful in their task, they are willing to enjoy the tool that is providing them success. Although the context is different, we believe that if individuals perform well in the activity, then they will also have a higher satisfaction with the information format.

H5: Greater levels of effectiveness will have a positive effect on satisfaction.

H6: Greater levels of efficiency will have a positive effect on satisfaction.

Based on the hypotheses discussed in this section, the research model for this study is represented below in Figure 7.

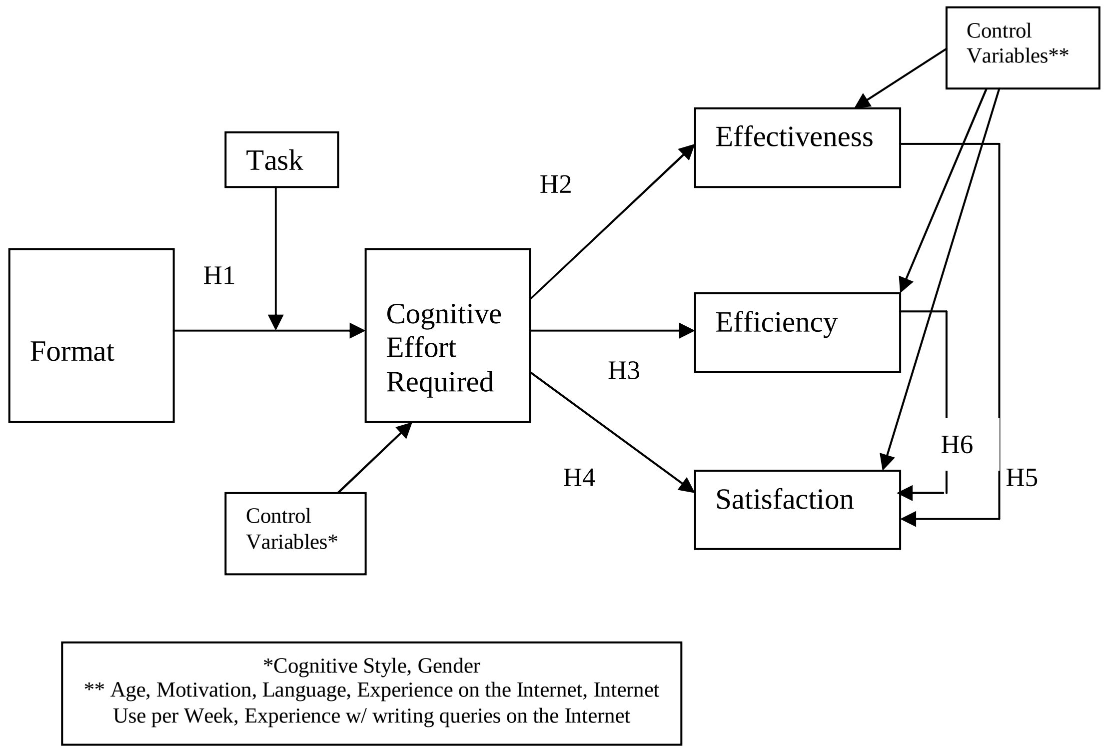


Figure 7. Research Model

6.4 Control variables

Demographic data such as motivation, age, gender, primary language, experience using search engines on the Internet, experience writing queries on search engines, and cognitive style will be collected for each participant for control purposes.

6.4.1 Motivation

Motivation is the amount of desire and willingness to complete the activity. Based on self-determination theory (Deci and Ryan, 2000), individuals are motivated intrinsically and extrinsically. Research shows that individuals with higher levels of motivation will perform better than those with lower levels of motivation. Motivation level will be collected to identify whether participants had a desire to truly participate in the study. Motivation will be used as a control variable.

We believe that an individual's motivation will affect his performance (effectiveness and efficiency) in completing the task regardless of whether there is a match between the format and task. Therefore, we will control for the level of motivation of participants on effectiveness, efficiency and satisfaction.

6.4.2 Age

Age is another factor that will be controlled for in this study. Older participants may take the experiment more seriously than younger participants, or older participants may not be as familiar with browsing information on the Internet as the younger participants. Therefore, we will control for age to ensure to account for its possible effects as a factor in determining effectiveness, efficiency and satisfaction.

6.4.3 Gender

Existing research shows that differences in gender may have an effect on technology. Bem (1981) shows that men and women process information using different socially constructed cognitive structures. For example, women make greater use of message cues in judging products (Meyers-Levy & Sternthal, 1991). It has also been found that women use a more comprehensive processing strategy whereas men tend to rely more on heuristics processing (Wollin, 2003; Darley & Smith, 1995). Therefore, we will control for gender to account for its potential influence on cognitive effort required.

6.4.4 Language

Participants whose primary language is not English may also have a more difficult time understanding and completing tasks, which could affect performance. Therefore, we will control for language in the study to account for its potential influence on effectiveness, efficiency and satisfaction.

6.4.5 Experience on the Internet and Writing Queries on Search Engines

Experience using Internet search engines may lead some participants to complete certain tasks faster than other participants who are unfamiliar with finding information on search engines. We also believe that individuals that have experience with writing queries for search engines may be more familiar with finding relevant information using search engines. Research has shown that education and training using information technology has positive effects on attitude and performance (Cheney et al. 1986). Bruner and Kumar (2000) found support for positive effects of Internet experience on attitudes toward web sites. Participants will be asked their experience of using Internet search engines and their experience with creating queries on search engines. These will be used as control variables to account for its potential influence on effectiveness, efficiency and satisfaction.

6.4.6 Cognitive Style

Since this study compares textual and graphical formats, we believe that an individual’s cognitive style may have an influence on how they may process and retrieve information on a particular type of format. Therefore, we will collect cognitive style as a control variable in this study. Cognitive style is defined as the individual way in which a person perceives, thinks, learns, solves problems and relates to others (Witkin et al. 1977). Hunt et al. (1989) define cognitive style as the way in which individuals' process and organize information, and arrive at judgments or conclusions based on their observations. In cognitive research, there are two different cognitive styles that have been identified. One type is described as analytical, deductive, rigorous, constrained, and formal (Nickerson et al. 1985). Another type is described as inductive, unconstrained, informal and creative (Nickerson et al. 1985). Allinson and Hayes (1996) call this the analysis-intuition dimensions (Refer to Table 6). Individuals with an analytic cognitive style tend to require facts and tend to be more interested in parts than a whole view of information (Allinson and Hayes 1996, Cools et al. 2006). They make judgments based on reason and focus on specific detail when processing information. We believe individuals with this type of cognitive style will process information better on a textual format; whereas, individuals with an intuitive cognitive style are more flexible and tend to be more interested in the whole rather than in parts (Allinson and Hayes 1996, Cools et al. 2006). They make judgments based on feelings and process information with a global approach. We believe individuals with this type of cognitive style will process information better on a graphical or visual format. Therefore, we will control for cognitive style on the effects of format on cognitive effort required.

For this study, we will measure cognitive style using the Cognitive Styles Index (CSI) developed by Allinson and Hayes (1996). CSI is a self-reported test designed to measure the analytic/intuitive dimension of cognitive style. The instrument contains 38 items, where individuals indicate a true/uncertain/false response. The test identifies an individual's cognitive style as being either analytical or intuitive. The CSI has a maximum score of 76. Higher scores toward 76 indicate a more intuitive cognitive style and lower scores toward zero indicate a more analytic cognitive style. Allinson and Hayes (1996) report test-retest reliability of the instrument at ($r = 0.90$, $p < 0.001$). The internal consistency measured by Cronbach's alpha range from 0.84 to 0.92.

Table 6. Description of the analytic–intuitive dimension (Cools et al. 2006)

Analytic pole	Intuitive pole
Convergence	Divergence
Sequential, structured	More randomly, less orderly
Facts, details	Possibilities, meanings, ideas
More interested in parts than in wholes	More interested in the whole than in the component parts
Logical, reflective	Impulsive, active
Conservative, conventional, conformity	Openness to experience, taking risks, subversive
Planned, organized, systematic	Flexible, spontaneous, open-ended
Utility	Novelty
Objective, impersonal, rational, intellectual	Subjective, (inter)personal, expressive
Verbal	Visual
Precision, methodicalness	Inventive, creative
Routine	Variety

6.5 Summary

Based on cognitive fit theory, we hypothesized how different formats fit different tasks that to reduce cognitive effort which would lead to higher performance and satisfaction. We also reviewed control variables that may have an effect on performance and satisfaction regardless of the match between format and task. In the next chapter, we review how the prototypes (formats) are created for this study.

CHAPTER 7

METHODOLOGY

7.1 Research Design

Two classic experiment designs are between subject designs and within subject (also referred to as repeated-measures design). Both experimental designs are discussed below to provide the rationale for why we used a mixed design experiment where there is a mixture of between subjects and within subject factors.

A repeated-measures research design uses a single sample of participants for each treatment condition. This design allows for researchers to control for differences between the participants since the same sample performs all the tasks for each treatment level. Second, when participants are difficult to recruit, repeated-measures designs are economical because each participant is measured under all conditions. However, there are disadvantages of carryover effects due to exposure to earlier levels in the treatment sequence. Participants could become familiar with the tasks as experience is gained. In addition, participants could also become tired with performing the tasks for each treatment and not focus on the latter tasks. In order to eliminate bias of gaining knowledge as the experiment progresses, participants can be assigned to different treatments in a distinctive sequence ("Between-Subject Versus Repeated-Measures Designs," n.d.).

In between subjects design, there are different participants for each level of treatment. For example, each participant would be exposed to one search engine, and comparisons would be made between the participants. The disadvantage to this design is that differences between subjects at each level of treatment could affect the results of the experiment. In order to eliminate some error that could arise from differences between subjects, participants will be randomized to one of the four treatment levels. Random assignment of participants to treatments ensures differences observed between subject groups are the result of the experimental intervention rather than differences between the subjects ("Between-Subject Versus Repeated-Measures Designs," n.d.). Another disadvantage in using this design is the large number of sample size needed to ensure proper statistical analysis. Since subjects only have to complete one treatment level rather than all treatment levels as in between-subject research design, the length of the experiment is shorter so that more subjects are willing to participate. In addition, participants do not become tired of performing the tasks. Therefore, this study used a mixed design incorporating both within subjects and between subject research designs.

There are four levels of treatments (standard textual, textual with categories, non-animated visualization, and animated visualization) with two groups for each treatment (Refer to Table 7). Participants were randomized to one treatment and perform each of the two tasks for that treatment. In each treatment, there were two groups. The order of open-ended and closed-ended tasks is switched for each group. This controls for any learning curve that could have developed by the participants while performing different tasks on the same prototype. Analysis was performed on initial data to ensure order effect does not exist. Subjects were asked to record their time as indicated on their survey sheet.

Table 7. Task Order and Groups for Levels of Treatment

Group	Assigned Format	Tasks Order
Group 1	Standard textual	Closed-ended Open-ended
Group 2	Standard textual	Open-ended Closed-ended
Group 3	Textual with categories	Closed-ended Open-ended
Group 4	Textual with categories	Open-ended Closed-ended
Group 5	Non-animated visualization	Closed-ended Open-ended
Group 6	Non-animated visualization	Open-ended Closed-ended
Group 7	Animated visualization	Closed-ended Open-ended
Group 8	Animated visualization	Open-ended Closed-ended

7.2 Protocol Analysis

We also performed protocol analysis, which is a methodology or technique that gathers verbal reports of thoughts and cognitive processes while completing a task (Owen et al. 2006, Mao and Benbasat 2000). This technique is rarely used in computing research; however, it is a useful methodology to extract expert knowledge, reasoning and experiences to determine decisions and ideas (Owen et al. 2006). Protocol analysis collects data process information on why certain decisions were made and the reasoning for each act. Five participants from three treatment groups (standard textual, textual with categories and animated visualization) were given a recording device. These participants were asked to “think aloud” by verbally record their experience of using the tool and completing the tasks on the given format. “Think-aloud requires subjects verbalize their thoughts while performing a given task. The resulting stream of utterances help indicate the way a subject is reasoning about how to perform a task” (Owens et al. 2006, p 118).

The instructions given to these participants were as follows:

> "Say out loud every thought or reaction that passed through your mind as you interact with the tool and complete the task. For example, say out loud why you clicked on a particular circle/link or what are your feelings as you complete each task. It does not matter if your sentences are not complete, since you are not explaining to anyone. Just act as if you were alone in the room speaking loudly to yourself and express your thoughts as you complete the study." (Owens et al. 2006, Mao and Benbasat 2000)

Participants were also told if they are silent for more than 10 seconds, they would be reminded to keep talking. Ericsson and Simon (1993) found that thinking aloud has no significant effect on the quality of performance; therefore, we do not think that collecting process data decreased the performance level of subjects in the experiment.

This process data was collected to determine why the performance and satisfaction outcomes for the study were obtained. We are trying to identify whether the results of the study matches with what participants went through when completing the experiments. Through the process data, we were able to discover other issues that were not determined through the actual performance measures. We discovered how and why participants navigated through information on the search results page. In addition, we determined how they interpreted using the tool and the different cues. The data helped us identify the order of their actions to complete the tasks. By understanding their chosen order, we were able to understand how performance measures match whether the cues were a useful scent to help find relevant results. If subjects did not choose the most relevant clusters or circles first, we were able to determine that animation and labeling relevancy were not useful cues. We were also able to understand their feelings towards the tool. They were asked to verbalize their feelings and actions. Based on their statements, we determine their frustration with the tool or tasks. We also determined that subjects did not like the interface therefore; they didn't bother with interacting with the tool correctly or completing the tasks. The data was coded to remove noise or irrelevant information. The process of coding involved mapping the data to the performance measures. The coding process is further described in Section 8.6.

We selected five subjects in each of the four treatments. We selected standard textual interface as a control to analyze the data collected from the other three treatments.

7.3 Development of task questions

There are a series of questions that comprise the open-ended and closed-ended tasks (Refer to Table 8). These questions were selected based on characteristics of questions asked in similar studies (Turetken and Sharda, 2005; Chung et al., 2005). The tasks will be selected from a broad range of topics, including sports, movies, news, literature, and automotive and local news. Closed-ended tasks involve a specific objective and finding accurate answers to the task. For example, an MSN Live or Google search for the query "who won the Nobel Prize in Medicine in 1979?" would produce a specific answer. In closed-ended tasks, individuals tend to find results rather quickly since the top results from a query on a search engine are likely to have the answers. Open-ended tasks involve finding relevant information on a general topic. Individuals may have to browse many search results to fully understand the topic and make a valid judgment to complete the task. Each question will have relevant information available in the returned results page, as judged by three experts ahead of time and verified by participants during the experiment. The experts are from academia and industry. One expert will be from the academic environment and two experts will be from the professional environment. These experts will be selected based on their knowledge and experience in technology. They will complete the activity with the same results that will be provided to the participants in the study. The answers provided by the experts on all tasks will be combined to serve as an "answer sheet" to judge the answers of participants completing the study.

Table 8. Sample size for Small, Medium and Large Effect size

Closed-ended Tasks
Question 1: What is the title of the article written by Ronen Feldman on text mining approaches?
Question 2: Which University is the 2008 Top National University in the US as indicated by USNews?
Question 3: Name the new video series that was created by the Official Visitor Site Greater Philadelphia organization to explore 24 of Philadelphia's finest neighborhoods.
Question 4: In what year was Temple University established?
Question 5: Who was the 32nd US President?
Open-ended Tasks
Question 6: Describe the similarities and differences in the political views in education of Hilary Clinton, Barack Obama, and John McCain.
Question 7: Imagine you are writing a paper on music. Describe how the genre of "hip hop" in 2008 has influenced our culture.

7.4 Pilot Studies

Two series of pilot studies were preformed prior to starting the actual experiment. In each pilot various changes were made in order to ensure a robust survey and tool were created for testing.

7.4.1 Pilot 1

In early Spring 2008, a pilot was conducted on 20 students. The experiment took place in the Biztech Lab (Lab 29) in Speakman Hall. The tools were created on Server 1. However, during the experiment, participants had a difficult time with the visualization prototypes. The pages were taking more than a few minutes to load which affects the performance of the tool. In addition, the survey questions, twelve in all, were listed on hard copies where participants had to fill in the results along with the website that they found their results. This was time consuming and took students about 2 hours to complete the experiment.

Changes made after Pilot 1:

- Moved all files, applications and interfaces from Server 1 to Server 2 for faster performance of the prototypes
- Created training instructions on web pages for participants to view at any time
- Created training questions for subjects to work to become comfortable with using the prototype
- Decreased the number of question from 12 to 10

- Create online survey instead of hard copy of survey (to help with copying the web site's URL on an online survey rather than writing the web site's URL on paper)

7.4.2 Pilot 2

In late Spring 2008, another pilot was conducted after the changes from Pilot 1 were incorporated into the study. The visualization prototypes did not take long to open and load. There was significant improvement in speed and performance of the prototypes. Participants were informed of how to use the tool and felt the instructions and training on the tool was helpful. Participants seemed pleased to use the online survey to copy websites into text boxes to answer questions. However, the length of the experiment was still taking approximately 90 minutes to complete.

Changes Made after Pilot 2 and Committee Feedback:

- Reduced the number of questions from 10 to 7
- Changed the types of questions asked that might be interesting to participants
- A cognitive style instrument is incorporated into the survey
- For the clustered format, the number of web sites within each category is added into the format
- Incorporated a method for protocol analysis for data collection

7.5 Sample Size

This study will require approximately 179 participants. Previous studies with similar research studies found significant results using 30 participants in an eight-cell research design (Chung et al., 2005) or 78 participants in a nine-cell research design (Turetken and Sharda, 2005). In this study, we used an application called G*Power to determine a sample size. The sample size is computed as a function of power level, significance level and population effect size in the population. The power of a study is its ability to detect a difference in the outcome. The minimum acceptable power level to determine sample size set by convention is 80% (Walsh et al. 1999, De Stephano 2003). We assumed a significance level of .05% based on acceptable conventions stating that we accept that 5% of the results are due to chance (Walsh et al. 1999, De Stephano 2003). Since we have a 4x2 ANOVA model, the denominator df is 3 [(4-1) (2-1)] and the number of treatment groups in the study is 8 (4x2). Table 9 provides estimates of the sample size for large, medium and small effect sizes. Since it is impractical to recruit 1095 participants to detect a small effect size, we will attempt to recruit 179 participants for detection of a medium effect size.

Table 9. Sample size for Small, Medium and Large Effect size

	Small Effect Size	Medium Effect Size	Large Effect Size
Effect size f	.40	.25	.10
α err prob	.05	.05	.05
Power (1-β err prob)	.80	.80	.80
Numerator df (degrees of freedom)	3	3	3
Number of treatment groups	8	8	8
Total sample size	1095	179	73

7.6 Experimental Procedure

Prior to starting the experiment, the researcher discussed the background of the study and their goal in completing the experiment. Participants will then be given training through written instructions on the web page for their treatment group on how to use the interface. To become familiar with using their assigned tool, participants will complete a sample question. Participants will be instructed to raise their hand if they have any questions during the training or the experiment so that a facilitator can help them. For participants in the protocol analysis methodology, there will be further instructions provided as stated earlier in the addendum.

Participants completed the training and initial surveys before starting the experiment to complete the task. The entire experiment will take participants approximately 90 minutes to complete. The time for each part of the experiment is outlined in Appendix D.

7.7 Summary

In this chapter, we discussed the research methodology to test the hypotheses proposed for this study. We discussed the process for the experiments to collect quantitative data and the process for protocol analysis to collect qualitative data. We reviewed the design of tasks for the experiment, the sample size and experimental procedure for the study. We also looked at the results of the two pilot studies that were conducted to update the prototypes (formats) and procedure for the actual experiments.

CHAPTER 8

DATA ANALYSIS

The purpose of this study was to study the effects of visualization and animation to reduce cognitive effort required to complete a particular task which in turn would increase performance and satisfaction. This chapter will present findings from statistical analyses of the data obtained in this study. This chapter begins with demographic information regarding the sample used including age, gender, language, motivation, cognitive style and experience using and writing queries on the Internet. The findings are then organized to match the two part analysis of the model. A mixed model ANOVA was conducted to examine the effects of cognitive effort required for closed and open ended task on a given format. A multivariate regression using a macro written by Preacher and Hayes (2008) was also conducted to assess the performance and satisfaction measures (see Figure 8). Statistical analyses were calculated using Statistical Package for Social Sciences (SPSS) software version 16.0.

Qualitative analysis was also conducted to identify why some of the results were obtained. Through the qualitative analysis, we are able to provide a more complete understanding of the results and relationships found in the statistical analysis. We are also able to identify other issues that may not be determined through the statistical measures, such as better understand which cues were useful scents for subjects to help find relevant search results.

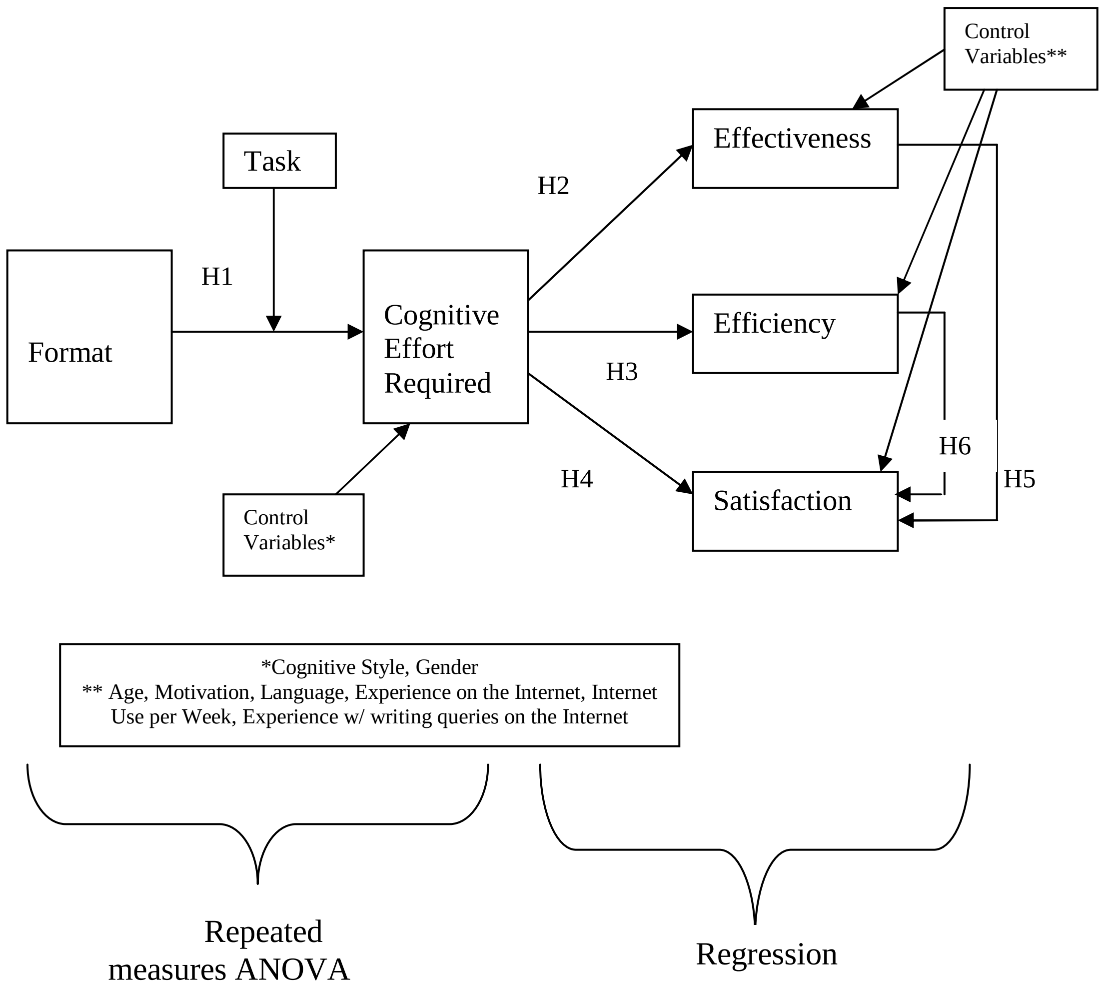


Figure 8. Research model labeled with statistical approaches

8.1 Description of participants

Students in both undergraduate and graduate programs were recruited at Temple University (n = 181). Professors were asked to provide students with extra credit for participating in the study. Since not all the professors provided extra credit, we had other incentive plans to recruit students. Additional participants from Accenture were also recruited to participate (n = 16); however the number of employees that volunteered was minimal. We analyzed the data from academia and industry together (total = 197) because there is no reason to believe that there is a difference between the two groups in completing the tasks. The questions were related to general searching and browsing tasks on the Internet. In addition, various variables (age, experience using the Internet and writing queries on the Internet) were controlled to identify any possible differences that would be related to this study.

Our incentives for participation in this study included either raffle to win a prize (iPod shuffle) or extra credit in the participant's class. Additional incentives of gift certificates were provided based on performance in the experiment and gift. All participants were eligible for gift certificates based on scores. In each treatment group, three participants with the highest score in effectiveness in both tasks will be awarded gift certificates ranging from 20 to 50 dollars (1^{st} prize will be \$50, 2^{nd} prize will be \$30 and 3^{rd} prize will be \$20). We believe that this type of incentive on performance will enable participants to perform well on the tasks rather than participate just for money.

Demographic data such as age, gender, cognitive style, motivation, experience using the Internet, and primary language was collected as described in the previous section for each participant for control purposes.

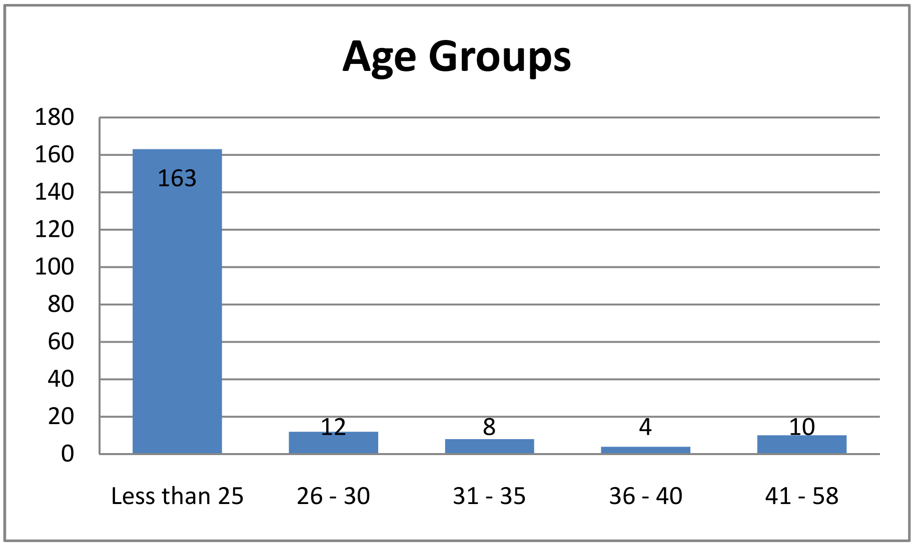


Figure 9. Age of Subjects

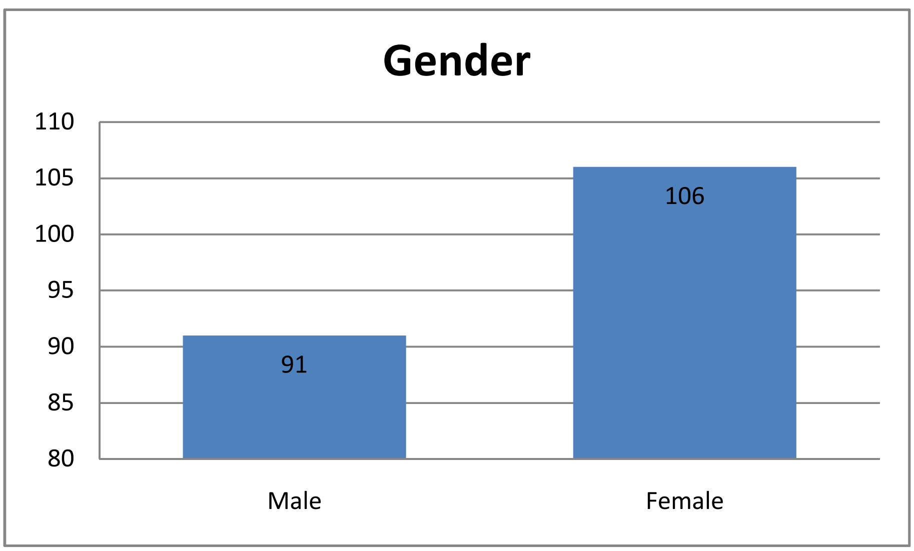


Figure 10. Gender of Subjects

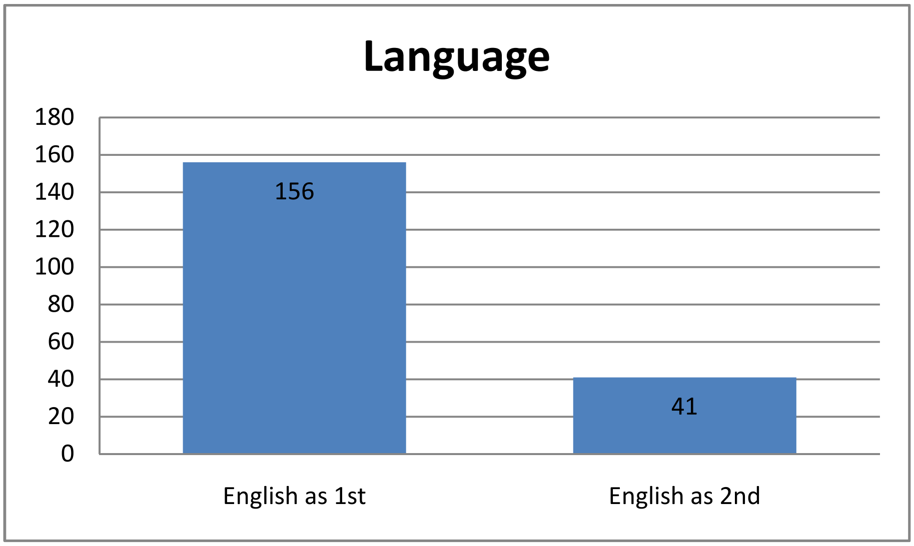


Figure 11. Frequency of English as a First Language among Subjects

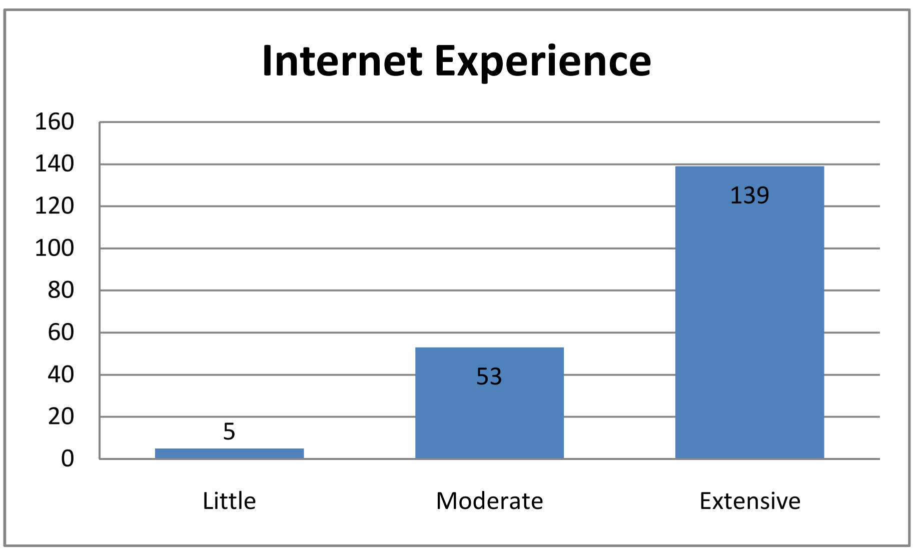


Figure 12. Level of Internet Experience of Subjects

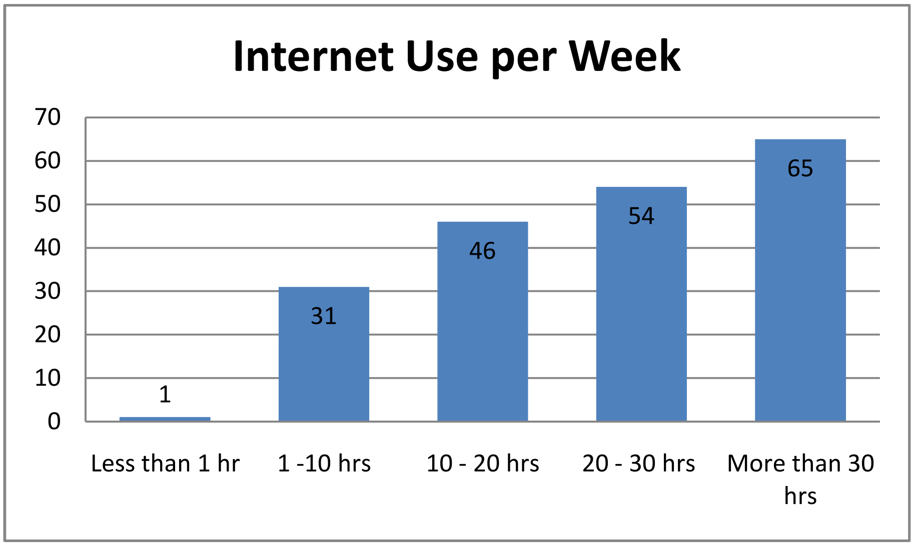


Figure 13. Level of Internet Use per Week of Subjects

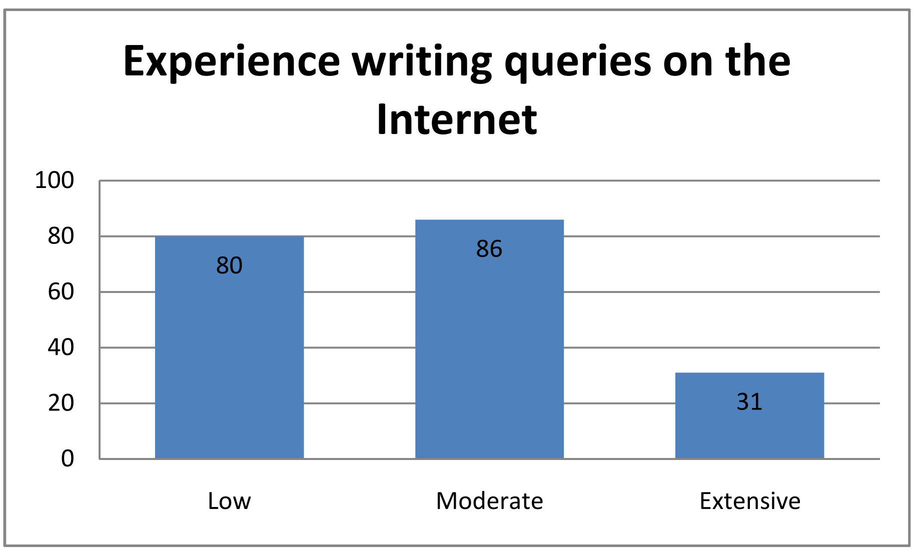


Figure 14. Level of Experience Writing Web Search Queries by Subjects

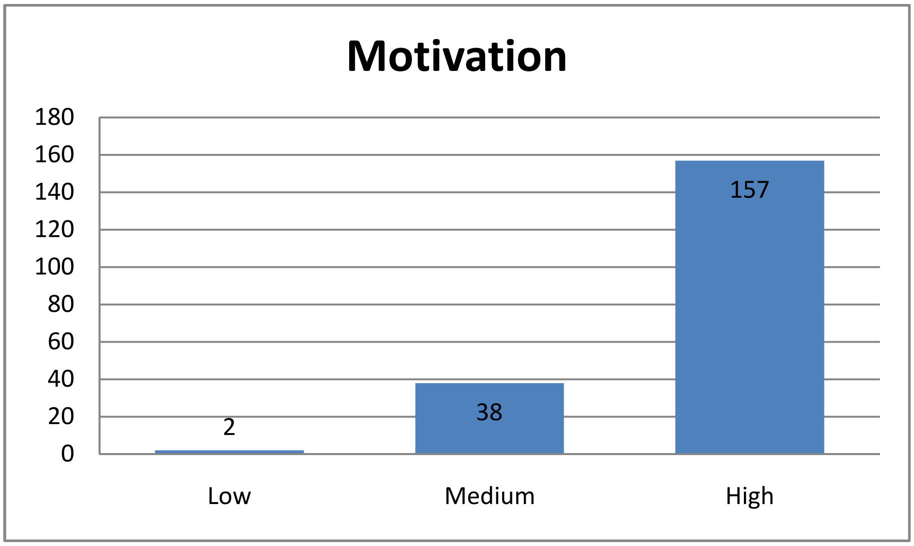


Figure 15. Self-assessed Motivation Levels of Subjects

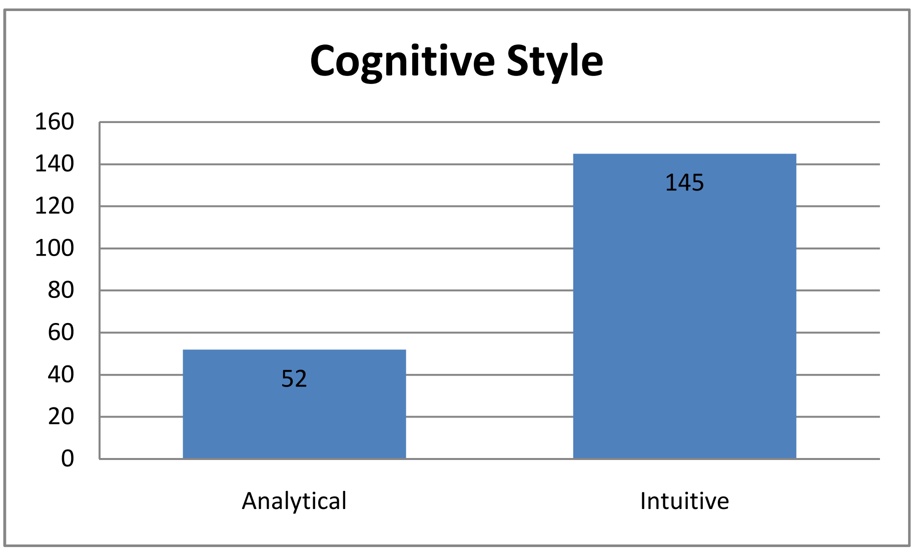


Figure 16. Cognitive Style of Subjects

8.2 Validity of measures

An instrument was administered to all participants to determine their cognitive style. The test identifies an individual's cognitive style as being either analytical or intuitive (see section 4.4.6 Cognitive Style for definitions). The CSI has a maximum score of 76. Higher scores toward 76 indicate a more intuitive cognitive style and lower scores toward zero indicate a more analytic cognitive style. In the sample size, we see there were 145 participants with intuitive style indicating that they would favor a spatial or visualization format compared to the 52 participants that were analytical favoring a more standard format.

The items used for measuring the constructs were taken from measures used in prior studies and some items were modified for this research context. To confirm the validity of the measures, convergent and discriminate validities will be evaluated on the measures for motivation, satisfaction and cognitive effort using principal component factor analysis. Convergent validity provides evidence that items used in the survey that should be related to each other based on theory are related, whereas discriminate validity provides evidence that items in the survey that should not be related to each other are not related.

Principal factor loading was conducted separately and jointly on motivation, satisfaction, and cognitive effort. Table 10 provides the loadings for each construct. The Kaiser-Meyer-Oklin (KMO) measure was used, which states if two variables share a common factor with other variables, their partial correlations will be small. Therefore, the closer KMO is to 1.0, the better the degree of commonality between the variables.

Generally, values above 0.50 are acceptable. The KMO for the variables in the survey are provided in Table 10, which show that all the variables are acceptable.

Table 10. KMO values for variables in survey

Variable	KMO
Motivation	.718
Satisfaction	.870
Cognitive Style	.666
Cognitive Effort	.911

Cronbach's alpha, which is a coefficient of reliability, was determined for motivation, satisfaction, cognitive style and cognitive effort. Cronbach's alpha measures how well a set of variables measures a single construct. When data have a multidimensional structure, Cronbach's alpha will usually be low. Values of 0.70 and higher are acceptable measures for reliability. Motivation, satisfaction, cognitive style and cognitive effort all have Cronbach's alpha higher than the acceptable value. See Table 11 below.

Table 11. Cronbach's alpha for variables in survey

Variable	Cronbach's alpha
Motivation	.775
Satisfaction	.914
Cognitive Style	.772
Cognitive Effort	.922

8.3 Testing Cognitive Fit

The hypotheses regarding cognitive fit are tested in this section. See Table 12.

Table 12. Hypotheses for Cognitive Fit between Format and Task

Closed-ended Tasks	H1a: there will be less cognitive effort required when using a textual format over a visualization format
	H1b: there will be less cognitive effort required when using a standard textual format over a clustered textual format
	H1c: there will be less cognitive effort required when using anon-animated visualization format over an animated visualization format
Open-ended Tasks	H1d: there will be less cognitive effort required when using a visualization format over a textual format
	H1e: there will be less cognitive effort required when using a clustered textual format over a standard textual format
	H1f: there will be less cognitive effort required when using an animated visualization format over a non-animated visualization format

The raw was tested for normality. All tests passed to verify that the data is normal (Refer to Appendix F to view the data and plots). To examine the differences in cognitive effort required of individuals using a particular format for different tasks, a repeated measure for mixed factorial ANOVA model was performed (see Figure 8). A mixed factorial model incorporates both within-subjects and between-subject data. The cognitive effort measure for closed and open tasks is the 2-level factor for within subject variables. Since participants had to complete both closed and open tasks on the given format, the results for cognitive effort on both of tasks are provided by the same participant this variable is entered in within subjects. The format variable is the between subject factor since there were different participants for each format.

The test of within-subjects shows the interaction between task (CETask: closed and open) and format (GroupID) is statistically significant at the 1% level: $F = 4.187$; $p = 0.007$ (See Table 13) which implies there is a difference in cognitive effort of tasks across formats. Since there is an interaction, the test of between-subject effects shows the effect of format (GroupID) on cognitive effort (CETask) is statistically significant at the 1% level: $F = 5.914$; $p<0.01$ (See Table 14).

Table 13. Test for Within-Subjects Contrasts for Cognitive Effort of Tasks and Format (GroupID)

Measure:Cognitive Effort

Source	CETask	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
CETask	Linear	2.137	1	2.137	1.525	.218	.008
CETask * cognitivestyle	Linear	.460	1	.460	.328	.567	.002
CETask * Sex	Linear	.563	1	.563	.402	.527	.002
CETask * GroupID	Linear	17.598	3	5.866	4.187	.007	.062
Error(CETask)	Linear	267.580	191	1.401			

Table 14. Repeated Measures Analyses for Test of Between Subjects of Format (GroupID) on Cognitive Effort (CETask) for Closed and Open Tasks

Measure: CETask
Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	1539.985	1	1539.985	494.548	.000	.721
Cognitive Style	30.384	1	30.384	9.757	.002	.049
Sex	.045	1	.045	.014	.905	.000
GroupID	55.249	3	18.416	5.914	.001	.085
Error	594.760	191	3.114			

Since the interaction between format (GroupID) and task (CETask: closed and open) is statistically significant ($p = 0.007$); an ANOVA was conducted separately on closed-ended and open-ended tasks to determine the significance of the effects of cognitive effort on each task. The control variables cognitive style and sex were included in the analyses. Sex is not significant ($p = .905$); however, cognitive style is significant ($p = .002$). The interaction effects of the control variables and format (GroupID) were not significant.

Hypotheses, H1a and H1d, involved comparing the two textual groups (standard format and clustered format) with the two visualization groups (non-animated format and animated format). Therefore, the data set involved adding a column to concatenated standard group (Format 1) and clustered group (Format 2) into Group 1 and non-animated group (Format 3) and animated group (Format 4) into Group 2. The effect of the concatenated formats on cognitive effort for closed-ended tasks is significant at the

1% level; F = 19.699; p = 0.000. Group 1 (standard format and clustered format) had a lower mean cognitive effort (3.714) than Group 2 (non-animated and animated format), which had a mean cognitive effort of 4.605. Therefore, H1a is supported, which states that users will experience a lower cognitive effort for textual formats over visualization formats. The same ANOVA test was conducted using the concatenated groups and open-ended tasks. The effect of the concatenated formats on cognitive effort for open-ended tasks is not significant (F = 1.695; p = .194).Therefore, H1d is not supported. See Appendix G and H for ANOVA output for the analyses for concatenated formats on cognitive effort.

In order to test hypotheses H1b, H1c, H1e, H1f, an ANOVA test was conducted; however instead using the concatenated format groups as the independent variable, the actual format groups were used. The effect of format on cognitive effort for closed-ended tasks is statistically significant (See Table 15) at the 1% level; F = 9.573; p = 0.000. The control variable of cognitive style is significant on cognitive effort for closed-ended tasks (p = 0.014); whereas, sex is not significant (p = 0.621).

Table 15. ANOVA results for Cognitive Effort for Closed-ended Tasks (CEclosed) by GroupID (Format)

Dependent Variable:CEclosed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	63.975[a]	5	12.795	6.793	.000	.151
Intercept	828.429	1	828.429	439.786	.000	.697
Sex	.463	1	.463	.246	.621	.001
Cognitive Style	11.684	1	11.684	6.202	.014	.031
GroupID	54.099	3	18.033	9.573	.000	.131
Error	359.788	191	1.884			
Total	3830.305	197				
Corrected Total	423.764	196				

a. R Squared = .151 (Adjusted R Squared = .129)

The pairwise comparisons between each format (GroupID) for cognitive effort of closed-ended tasks are provided in Table 16. The standard format group did better for closed-ended tasks compared to all formats and difference between each group is significant. There is a partially significant mean difference between standard format group and clustered format group (M= -.700; $p<0.10$), between standard format group and non-animated format group (M= -1.042; p = 0.002), between standard format group and animated format group (M = -1.451; p = 0.00). For the clustered format group, there is only a statistically significant mean difference when compared to the animated visualization format group (M = .281; p =0.05). For non-animated format group, there is only a statistically significant mean difference with the standard format group (M = 1.042;

p = 0.002) and for animated format group, there is only a statistically significant mean difference with the standard format group (M = 1.451; p = 0.000) and the clustered format group (M = .751; p = .050). Therefore, hypotheses H1b and H1c, are not supported.

Table 16. Pairwise Comparisons for Cognitive Effort for Closed-ended Tasks (CEclosed) by Formats (GroupID)

Dependent Variable:CEclosed

(I) GroupID	(J) GroupID	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Standard	Clustered	-.700	.278	.075	-1.441	.040
	Non-animated	-1.042*	.280	.002	-1.789	-.296
	Animated	-1.451*	.282	.000	-2.202	-.700
Clustered	Standard	.700	.278	.075	-.040	1.441
	Non-animated	-.342	.274	1.000	-1.071	.388
	Animated	-.751*	.281	.050	-1.501	.000
Non-animated	Standard	1.042*	.280	.002	.296	1.789
	Clustered	.342	.274	1.000	-.388	1.071
	Animated	-.409	.283	.900	-1.164	.346
Animated	Standard	1.451*	.282	.000	.700	2.202
	Clustered	.751*	.281	.050	.000	1.501
	Non-animated	.409	.283	.900	-.346	1.164

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

For open-ended tasks, the effect of format (GroupID) on cognitive effort (CEopen) is not statistically significantly at the 1% level; F = 2.375; p = 0.071. The control variable of cognitive style is significant on cognitive effort for open-ended tasks (p = 0.008); whereas, sex is not significant (p = 0.815). See Table 17 below for results.

Table 17. ANOVA Results for Cognitive Effort of Open-ended Tasks (CEopen) by Format (GroupID)

Dependent Variable:CEopen

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	32.719[a]	5	6.544	2.487	.033	.061
Intercept	713.693	1	713.693	271.246	.000	.587
Sex	.145	1	.145	.055	.815	.000
cognitivestyle	19.160	1	19.160	7.282	.008	.037
GroupID	18.748	3	6.249	2.375	.071	.036
Error	502.551	191	2.631			
Total	3129.595	197				
Corrected Total	535.270	196				

a. R Squared = .061 (Adjusted R Squared = .037)

The pairwise comparisons between each format (GroupID) for cognitive effort of open-ended tasks (CEopen) are provided in Table 18. There is no statistically significant difference between each format group as seen in Table 18. Therefore, there is no support for hypotheses, H1e and H1f.

Although there are no statistically significant differences between the groups, we provide the mean cognitive effort for each format (GroupID) to understand the direction of cognitive effort required on each group for closed-ended and open-ended tasks. On the standard format, the mean cognitive effort is less for closed ended tasks (M = 3.3333, SD = 1.34) than for open-ended tasks (M = 3.4885, SD = 1.49). On the clustered format, the mean cognitive effort is greater for closed ended tasks (M = 4. 1298, SD = 1.39) than for open ended tasks (M = 3.5125, SD = 1.48). On the non-animated format, the mean cognitive effort is greater for closed ended tasks (M = 4.440, SD = 1.61) than for open-ended tasks (M = 3.4480, SD = 2.01). On the animated format, the mean cognitive effort is greater for closed-ended tasks (M = 4.7271, SD = 1.18) than open-ended tasks (M = 4.0802, SD = 1.54). Figure 17 below provides the mean cognitive effort for closed-ended and open-ended tasks for each format. In Figure 17, we see that for closed ended tasks, the animated visualization format have a higher cognitive effort over the other formats, next was the non-animated, then clustered and last standard format, which had the least cognitive effort required.

Table 18. Pairwise Comparisons of Cognitive Effort for Open-ended Tasks (CEopen) by Format (GroupID)

Dependent Variable:CEopen

(I) GroupID	(J) GroupID	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Standard	Clustered	.107	.328	1.000	-.768	.982
	Non-animated	.129	.331	1.000	-.754	1.011
	Animated	-.647	.333	.321	-1.534	.241
Clustered	Standard	-.107	.328	1.000	-.982	.768
	Non-animated	.021	.323	1.000	-.841	.884
	Animated	-.754	.333	.148	-1.641	.133
Non-animated	Standard	-.129	.331	1.000	-1.011	.754
	Clustered	-.021	.323	1.000	-.884	.841
	Animated	-.775	.335	.129	-1.667	.117
Animated	Standard	.647	.333	.321	-.241	1.534
	Clustered	.754	.333	.148	-.133	1.641
	Non-animated	.775	.335	.129	-.117	1.667

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

Although cognitive effort for open-ended tasks (CEopen) on format (GroupID) is not statistically significant, we see that for open-ended tasks, animated visualization has the highest cognitive effort required, then clustered format, standard format and non-animated format (See Figure 19). Therefore, for open-ended tasks, non-animated visualization have the least cognitive effort mean required; whereas, the animated visualization have the greatest cognitive effort mean compared to the rest of the formats for both closed-ended and open-ended tasks. For closed-ended tasks, the standard format has the least cognitive effort required.

Table 19. Mean of Cognitive Effort of Tasks (CEclosed and CEopen) for Formats

Format		**CEclosed**	CEopen
Standard	Mean	**3.3333**	3.4885
Clustered	Mean	**4.1298**	3.5125
Non-animated	Mean	**4.4398**	3.4480
Animated	Mean	**4.7271**	4.0802
Total	Mean	**4.1584**	3.6289
p-value		**0.000**	0.190*

*Not statistically significant

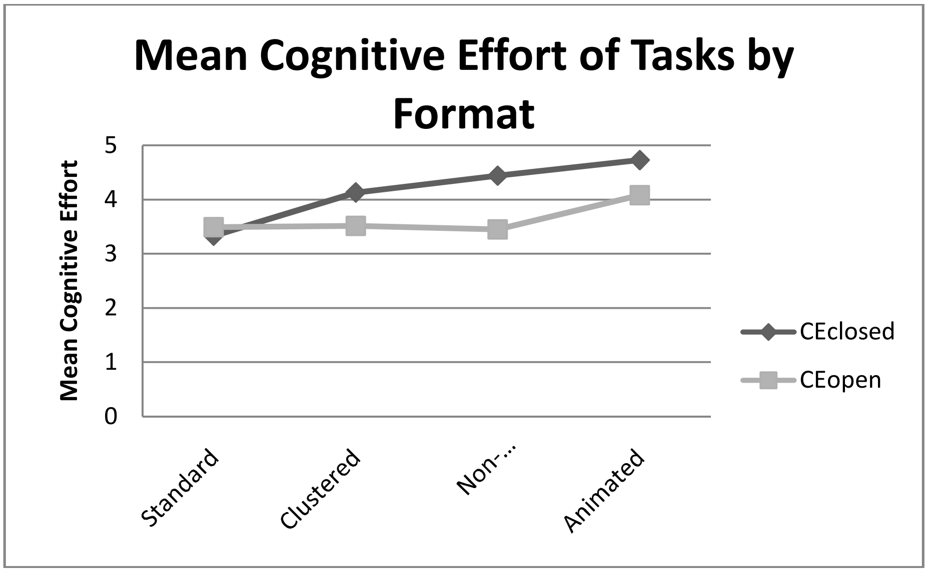


Figure 17. Estimated Marginal Means of Cognitive Effort for Formats by Tasks*

*(CEopen is not statistically significant)

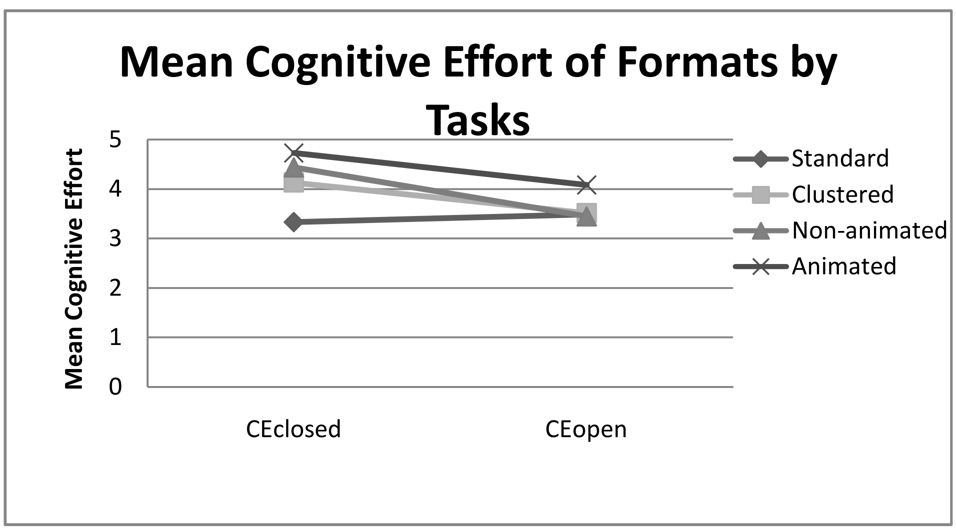


Figure 18. Estimated Marginal Means of Cognitive Effort for Formats by Tasks*

*(CEopen is not statistically significant)

Table 20. Summary of Cognitive Fit Hypotheses

Hypotheses	Supported or Not supported
H1a: For closed-ended tasks, there will be less cognitive effort required when using a textual format over a visualization format.	Supported
H1b: For closed-ended tasks, there will be less cognitive effort required when using a standard textual format over a clustered textual format.	Not Supported
H1c: For closed-ended tasks, there will be less cognitive effort required when using a non-animated visualization format over an animated visualization format.	Not Supported
H1d: For open-ended tasks, there will be less cognitive effort required when using a visualization format over a textual format.	Not Supported
H1e: For open-ended tasks, there will be less cognitive effort required when using a clustered textual format over a standard textual format.	Not Supported
H1f: For open-ended tasks, there will be less cognitive effort required when using an animated visualization format over a non-animated visualization format.	Not supported

8.4 Testing Performance and Satisfaction Measures

In this section, the hypotheses for performance and satisfaction are tested. Based on CFT, individuals will achieve a higher level of performance when there is lower cognitive effort based on a fit or match between information format and task. From our earlier discussion, we proposed that this type of fit enables individuals to perform better by making better decisions regarding which information is relevant to their search. We measure performance through effectiveness and efficiency (Chung et al., 2005; Turetken and Sharda, 2005; Roussinov and Chen, 2001; Vessey and Galletta, 1991). Satisfaction is measured to identify how pleased subjects were with format. From Figure 8, the variables of effectiveness and efficiency have an indirect effect on satisfaction or mediate the effect of satisfaction (Hypotheses H5 and H6). We believe that higher levels of performance will have a positive effect on satisfaction. Therefore, we need to account for the effects of effectiveness and efficiency on satisfaction in order to estimate the total effects of cognitive effort on satisfaction.

Table 21. Hypotheses for Performance and Satisfaction

Performance (Effectiveness)	H2: Effectiveness will be higher when there is less cognitive effort required.
Performance (Efficiency)	H3: Subjects will take less time to complete tasks when there is less cognitive effort required.
Satisfaction	H4: Satisfaction will be higher when there is less cognitive effort required.
Indirect Effects on Satisfaction (mediating effectiveness)	H5: Greater levels of effectiveness will have a positive effect on satisfaction.
Indirect Effects on Satisfaction (mediating efficiency)	H6: Greater levels of efficiency will have a positive effect on satisfaction.

To test the hypotheses on performance and satisfaction, regression analysis was conducted for both closed-ended and open-ended task together. Since our model includes effectiveness and efficiency as mediators for satisfaction, we also had to include the mediating effects, if any into the analysis. We tested the direct effects of cognitive effort on effectiveness, efficiency and satisfaction as well as the indirect effects of effectiveness and efficiency on satisfaction. The analysis was done through a macro written by Preacher and Hayes (2008) to utilize in SPSS. The macro tests direct and indirect effects of the entire model. In order to test the indirect effects on satisfaction, we must understand how to analyze mediating effects.

Based on analysis for mediating effects from Baron and Kenny (1986) (http://davidakenny.net/cm/mediate.htm#ST), various regression tests should be conducted between the independent, mediator and dependent variables to identify the influence of the mediator variables on the dependent variable. The first step is to determine the significance of the independent variable on the dependent variable (cognitive effort on satisfaction). The next step is to determine the significance of the mediator variable as an independent variable on the dependent variable (effectiveness and efficiency on satisfaction - separately). The third step is to determine the independent and mediator variables (as independent variable) on the dependent variable (cognitive effort, effectiveness and efficiency on satisfaction - separately). If the variables are significant in all the steps, then there is a mediating effect. The steps described are conducted below for closed-ended and open-ended tasks separately and for each mediator variable – effectiveness and efficiency.

The macro written by Preacher and Hayes (2008) generates "estimates for indirect effects in a multiple mediator model (Figure 19), where *c* is the *total effect* of *X* on *Y*, *c′* is the *direct effect* of *X* on *Y*, and the specific indirect effect of *X* on *Y* through mediator M_i is defined as $a_i b_i$" (Preacher and Hayes, 2008, http://www.comm.ohio-state.edu/ahayes/SPSS%20programs/indirect.htm). The macro reflects the steps outlined by Baron and Kenny (1986) to estimate the effects of mediation discussed earlier.

Effectiveness was computed using different calculations for closed-ended and open-ended tasks; therefore, the raw effectiveness measures for both tasks cannot be combined into one data set. In order to combine the data from closed-ended and open-ended tasks, the scores for effectiveness need to be standardized as z-scores for closed-ended and open-ended tasks (Vessey and Galletta, 1991). This ensures that the scores for effectiveness for each task are now measured using the same scale. The z-score is derived by subtracting the population mean from an individual raw score and then dividing the difference by the population standard deviation (Abdi, 2007). The z-score provides how high or low each effectiveness score was measured compared to the mean effectiveness for closed-ended and open-ended tasks separately. By standardizing the score for effectiveness for both tasks, we can combine the data set for closed-ended and open-ended tasks to perform statistical analyses. The values for effectiveness and satisfaction for closed-ended and open-ended tasks are measured on the same scale, so the z-scores for these variables do not need to be determined and the actual data set can be combined for both tasks. Once the data set for both tasks was combined, regression analyses using Preacher and Hayes's (2008) macro in SPSS was conducted.

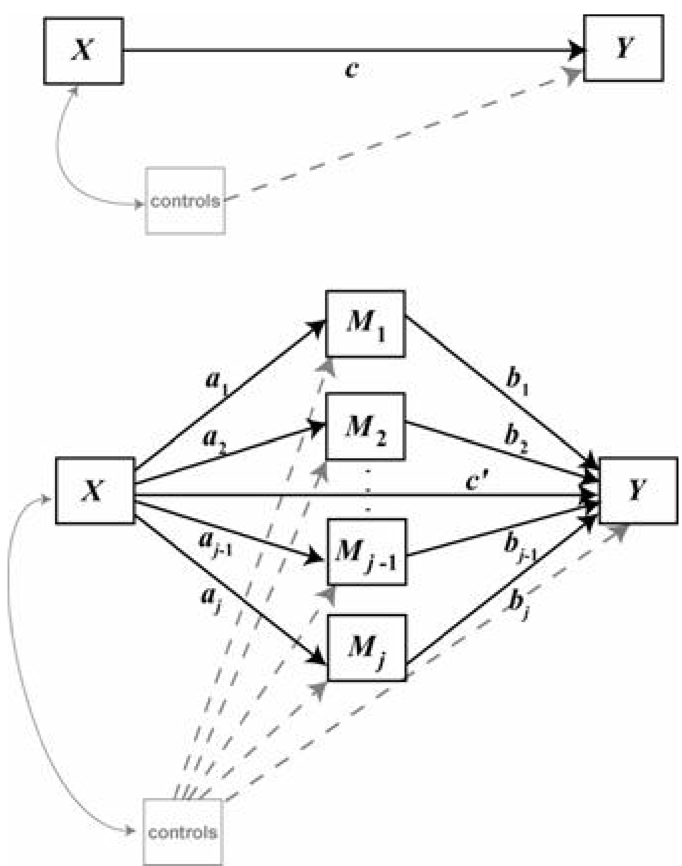


Figure 19 - A multiple mediator model

Retrieved from http://www.comm.ohio-state.edu/ahayes/SPSS%20programs/indirect.htm

The direct and indirect effects for the paths between cognitive effort (IV), effectiveness (M), efficiency (M) and satisfaction (DV) are provided in Table 22. Figure 20 provides the significance of the mediator variables on satisfaction based on the regression analyses. Although effectiveness and efficiency is a mediator for satisfaction, they are also dependent variables for a direct path from cognitive effort (IV). The macro output seen in Table 22 provides the direct paths for effectiveness (labeled as "a paths") and efficiency (labeled as "a paths") and satisfaction (labeled as "c` path"). According to Baron and Kenny (1986), we also need to test the mediators on the dependent variable, which is labeled as "b paths". The path for cognitive effort on satisfaction needs to account for the mediating variables, which is identified through "c path". In order for a mediating effect to occur, both "a path" and "b path" must be statistically significant. If either path "a" or path "b" is not significant, then there are no mediating effects.

From Table 22, we can see that cognitive effort required on efficiency (TIME) is not significant (p = .1549) and on effectiveness (Zscore) is significant (p = 0.0119). The mediator, effectiveness (Zscore) on satisfaction is significant (p = 0.000); however efficiency (TIME) is not significant (p = .7893). Since the path between the IV and M and M to the DV is significant for effectiveness, effectiveness has a mediating effect on satisfaction. The total effect of cognitive effort required including the effects of effectiveness on satisfaction is significant (p = 0.000).

The control variables of English (as a first language) (p = .0059) and experience on the Internet (p = 0.000) is significant on effectiveness, efficiency and satisfaction. When experience on the Internet increased (t = -4.1434), then performance and satisfaction increased. When English is not the first language (t = 2.7686), then performance and satisfaction decreased.

Table 22. Effects of Cognitive Effort Required on Effectiveness, Efficiency and Satisfaction (Combined Data Set of Closed-ended and Open-ended Tasks)

Pathway in Model	Coefficient	Standard Error	t-statistic	p-value
Cognitive Effort on Effectiveness (z-score)	-.0801	.0317	-2.5274	.012
Cognitive Effort on Efficiency	107.2369	74.2435	1.4252	.155
Effectiveness on Satisfaction	.3326	.0700	4.7520	.000
Efficiency on Satisfaction	.0000	.0000	-.2674	.789
Cognitive Effort on Satisfaction (w/ Mediators)	-.5518	.0447	-18.2839	.000
Cognitive Effort on Satisfaction (w/o Mediators)	-.5243	.0440	-17.6747	.000

Table 23. Significance of Control Variables for Combined Data Set

Control Variables	Coefficient	Standard Error	t-statistic	p-value
Age	-.0148	.0105	-1.4029	.1615
English as 1st Language	.4705	.1699	2.7686	.0059
Experience on the Internet	-.6038	.1457	-4.1434	.0000
Use of Internet per week	-.0943	.0687	-1.3721	.1708
Experience writing Queries	.0046	.1001	-.0456	.9637
Motivation Level	.0505	.0763	.6622	.5082

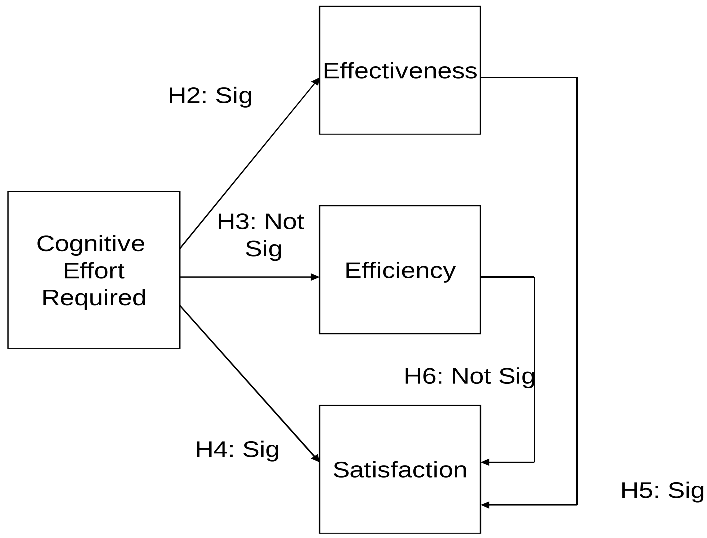


Figure 20. Significance of Variables for Combined Data (Both Tasks)

Table 24. Summary of Performance and Satisfaction Hypotheses for Combined Data Set

Hypotheses	Outcome
H2: Effectiveness will be higher when there is less cognitive effort required.	Supported
H3: Subjects will take less time to complete tasks when there is less cognitive effort required.	Not supported
H4: Satisfaction will be higher when there is less cognitive effort required.	Supported
H5: Greater levels of effectiveness will have a positive effect on satisfaction.	Not Supported
H6: Greater levels of efficiency will have a positive effect on satisfaction.	Supported

Since CFT states that there are higher performance outcomes depending on a fit between task and format, further analyses was performed to understand the effects of cognitive effort required on performance and satisfaction for closed-ended and open-ended tasks separately. The entire output provided from the macro is provided in Appendix I and J for closed-ended and open-ended tasks, respectively. The direct and indirect effects for the paths between cognitive effort (IV), effectiveness (M), efficiency (M) and satisfaction (DV) for closed-ended tasks are provided in Table 25. Figure 21 provides the significance of the mediator variables on satisfaction based on the regression analyses. The macro output seen in Table 25 provides the direct paths for effectiveness (labeled as “a paths”) and efficiency (labeled as “a paths”) and satisfaction (labeled as “c` path”).

Table 25. Direct Effects and Indirect Effects for Closed-ended Tasks

Pathway in Model	Coefficient	Standard Error	t-statistic	p-value
Cognitive Effort on Effectiveness	-.1927	.0523	-3.6856	.000
Cognitive Effort on Efficiency	87.1848	158.4772	.5501	.583
Effectiveness on Satisfaction	-.0043	.0637	-.0669	.947
Efficiency on Satisfaction	.0000	.0000	-1.9223	..056
Cognitive Effort on Satisfaction (w/ Mediators)	-.8406	.0460	-18.2839	.000
Cognitive Effort on Satisfaction (w/o Mediators)	-.8379	.0474	-17.6747	.000

Table 26. Significance of Control Variables for Closed-ended Tasks

Control Variables	Coefficient	Standard Error	t-statistic	p-value
Age	-.0157	.0103	-1.5264	.1286
English as 1^{st} Language	.2894	.1654	1.7494	.0819
Experience on the Internet	-.3897	.1422	-2.7408	.0067
Use of Internet per week	-.0954	.0670	-1.4247	.1559
Experience writing Queries	.0767	.0974	.7876	.4319
Motivation Level	.0609	.0742	.8201	.4132

The effects of cognitive effort (CEclosed) for closed-ended tasks is significant for effectiveness (SCOREclosed) at p = .0003 but not on efficiency (TIMEclosed) (p= .5829). Refer to Appendix I for the full output from the analyses. Therefore, the data analysis shows support H2 (effectiveness) and no support for H3 (efficiency) for closed-ended tasks. The t-values in the output in Table 25 show that when cognitive effort increased, then effectiveness decreases (t = -3.6856, p = 0.0003).

In order to estimate the effects of cognitive effort on satisfaction (H4), we must account for the mediating effects of effectiveness (H5) and efficiency (H6). Since the effect of cognitive effort on efficiency (a path) is not significant, there is no mediating effect of efficiency on satisfaction, regardless of the significance of efficiency on satisfaction (b path). Cognitive effort is statistically significant on effectiveness (a path); however, effectiveness is not significant on satisfaction (b path) as p = .9468. Therefore, there is no mediating effect of effectiveness on satisfaction. There is no support of H5 or H6 for closed-ended tasks.

The total effect of cognitive effort on satisfaction is significant at ($t = -17.9411$; $p = 0.0000$). This provides support for H4 (satisfaction). The t-values in Table 25 show that when cognitive effort increases, then satisfaction decreases.

The coefficients for the total effect of cognitive effort including the mediators ($B = -8406$) and the direct effect of just cognitive effort ($B = -8379$) are very close. This provides evidence that removing the mediator variables out of the model for closed-ended tasks does not change the coefficient drastically.

For closed-ended tasks, experience on the Internet is the only control variable that is significant ($t = -2.7408$; $p = .0067$). The results show that when experience on the Internet is high, then performance and satisfaction outcomes increases. We will discuss this further in the next chapter. As indicated in Table 26, none of the other control variables are significant. The interaction of the control variables with the independent variable also did not produce significant results.

The analysis was repeated for open-ended tasks. Refer to Appendix J for the full output of the analyses. The direct and indirect effects for the paths between cognitive effort (IV), effectiveness (M), efficiency (M) and satisfaction (DV) for open-ended tasks are provided in Table 27. Figure 22 provides the significance of the mediator variables on satisfaction based on the regression analyses. The effect of cognitive effort on effectiveness is not significant for open-ended tasks (SCOREopen) ($p = .7126$); therefore H2 for open-ended tasks is not supported. The effect of cognitive effort

(CEopen) for open-ended tasks on efficiency (TIMEopen) is significant at $p = 0.0027$; thus providing support for H3.

In order to estimate the effects on satisfaction, we need to understand how the mediating variables affect satisfaction. For open-ended tasks, cognitive effort on effectiveness is not significant (a path); therefore there is no mediating effect of effectiveness on satisfaction regardless of the significance of effectiveness on satisfaction (b path) ($p= 0.000$). There is no support for H5 for open-ended tasks. Since the effects of cognitive effort is significant on efficiency (a path) and the effects of efficiency is significant on satisfaction (b path) ($p= 0.000$), efficiency has a mediating effect on satisfaction. This supports H6 for open-ended tasks. The total effect of cognitive effort on satisfaction including the mediators is significant at $p = 0.0000$, which provides support for H4. The t-values show that when cognitive effort increases, time increase which implies efficiency decreases ($t = 3.0452$, $p = 0.0027$). When cognitive effort increases, then satisfaction decreases ($t = -4.1674$, $p = 0.000$).

For open-ended tasks, the control variable of experience on the Internet is significant ($t = -3.0997$; $p = 0.0022$). The results show that when experience on the Internet increases, then performance and satisfaction outcomes increases. As indicated in Table 28, none of the other control variables are significant. The interaction of the control variables with the independent variable also did not produce significant results.

Table 27. Direct and Indirect Effects for Open-ended Tasks

Pathway in Model	Coefficient	Standard Error	t-statistic	p-value
Cognitive Effort on Effectiveness (open)	-.0017	.0047	-.3689	.7126
Cognitive Effort on Efficiency (open)	161.0958	52.9013	3.0452	.0027
Effectiveness on Satisfaction (open)	4.5854	1.0333	4.4377	.0000
Efficiency on Satisfaction (open)	.0005	.0001	4.9595	.0000
Cognitive Effort on Satisfaction (w/ Mediators) (open)	-.3069	.0737	-4.1674	.0000
Cognitive Effort on Satisfaction (w/o Mediators) (open)	-.3723	.0683	-5.4482	.0000

Table 28. Significance of Control Variables for Open-ended Tasks

Control Variables	Coefficient	Standard Error	t-statistic	p-value
Age	-.0242	.0168	-1.4350	.1530
English as 1^{st} Language	.6538	.2746	2.3807	.0183
Experience on the Internet	-.7247	.2338	-3.0997	.0022
Use of Internet per week	-.1170	.1105	-1.0593	.2908
Experience writing Queries	.0449	.1610	.2791	.7805
Motivation Level	.0553	.1228	.4505	.6529

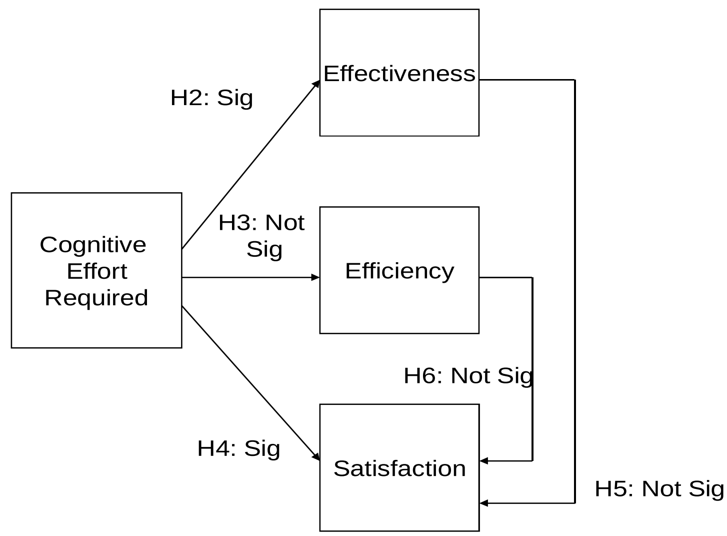


Figure 21. Significance of Variables for Closed-ended tasks

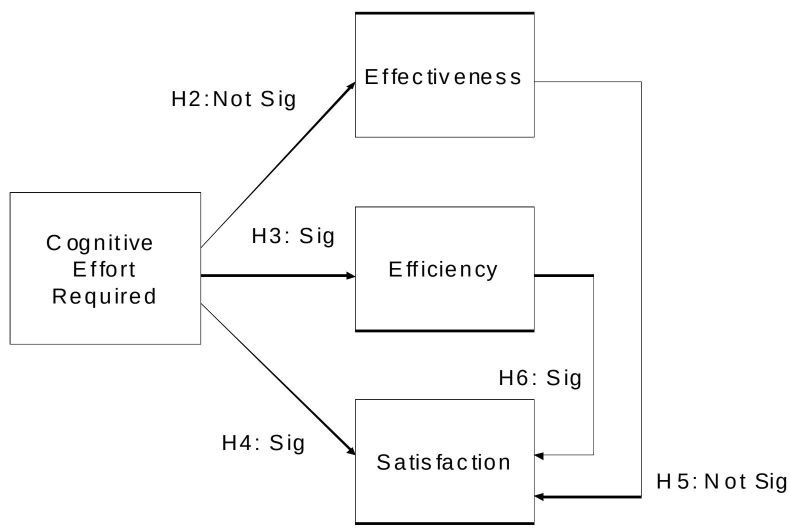


Figure 22. Significance of Variables for Open-ended tasks

Table 29. Summary of Performance and Satisfaction Hypotheses for Closed-ended and Open-ended Tasks (separate data sets)

Hypotheses	Closed-ended Tasks	Open-ended Tasks
H2: Effectiveness will be higher when there is less cognitive effort required.	Supported	Not Supported
H3: Subjects will take less time to complete tasks when there is less cognitive effort required.	Not supported	Supported
H4: Satisfaction will be higher when there is less cognitive effort required.	Supported	Supported
H5: Greater levels of effectiveness will have a positive effect on satisfaction.	Not Supported	Not Supported
H6: Greater levels of efficiency will have a positive effect on satisfaction.	Not Supported	Supported

8.5 Additional Statistical Analyses

Although effectiveness is statistically significant for closed-ended tasks and not for open-ended tasks, we look at the effectiveness for each task across the different formats to determine the direction of effectiveness across tasks. We performed ANOVA on each dependent variable separately using format (GroupID) as the independent variable. For closed-ended tasks, the standard format has a greater mean effectiveness, followed by clustered, non-animated and last animated. For open-ended tasks, the standard format had a greater mean effectiveness followed by clustered, animated and last non-animated formats. Figure 23 below provides the mean effectiveness for closed-ended and open-ended tasks for each format. In Figure 23, we see that for both closed ended and open-ended tasks, the standard format has a higher effectiveness over the other formats.

The effects of efficiency is statistically significant for open-ended tasks but not for closed-ended tasks. However, we look at efficiency for each format to understand the direction of efficiency across the tasks. On the standard format groups, the mean efficiency was similar for both closed-ended (M = 15:17, SD = 07:21) and open-ended tasks (M = 15:13, SD = 10:49). For clustered format groups, efficiency was greater for open-ended tasks (M = 19:38, SD = 08:08) compared to closed-ended tasks (M = 12:11, SD = 07:04). For non-animated formats, efficiency was greater for open-ended tasks (M = 25:46, SD = 15:47) than for closed-ended tasks (M = 08:31, SD= 35:52). For animated format groups, efficiency was greater for open-ended tasks (M = 14:20, SD = 17:39) than for closed-ended tasks (M = 38:18, SD = 1:44:49). Therefore, efficiency was greater (less time) to complete open-ended tasks using clustered, non-animated and animated formats. In Figure 24, for open-ended tasks, efficiency was greater for non-animated format group, followed by clustered format group, then animated format group and finally standard format group. It seems that the non-animated visualization format groups and clustered format groups were labeled with relevancy which could explain the greater efficiency in completing the tasks.

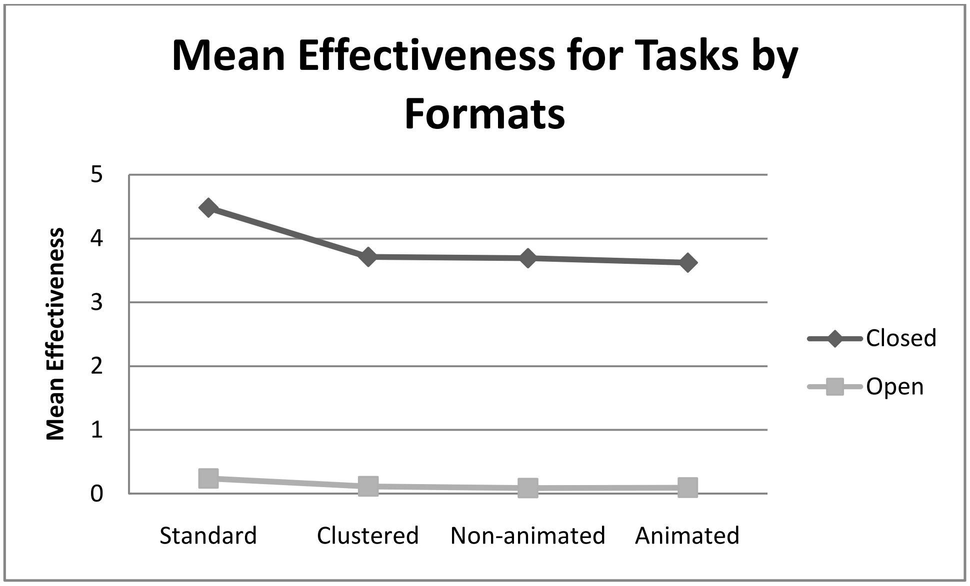


Figure 23. Estimated Marginal Mean Effectiveness for Tasks by Formats*

*Differences in effectiveness across groups are not statistically significant for the open-ended task

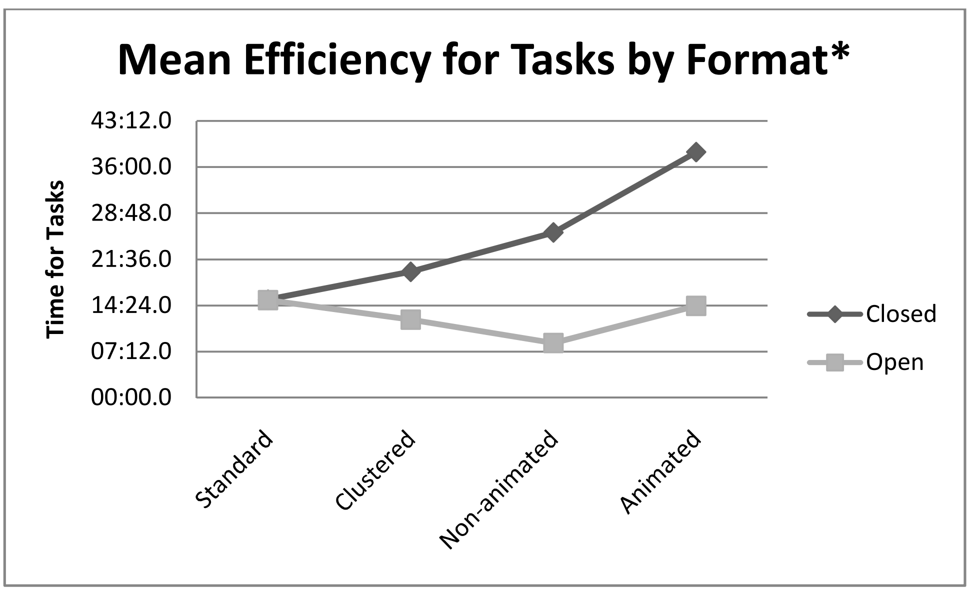


Figure 24. Estimated Marginal Mean Efficiency for Tasks by Formats*

*Differences in efficiency across groups are not statistically significant for the closed-ended task

The effect of satisfaction is significant for both closed-ended and open-ended tasks. On the standard format, the mean satisfaction is higher for closed-ended tasks ($M = 4.562$, $SD = .220$) than for open-ended tasks ($M = 4.212$, $SD = .255$). On the clustered format, the mean satisfaction is higher for open-ended tasks ($M = 4.215$, $SD = .245$) than for closed-ended tasks ($M = 3.923$, $SD = .211$). On the non-animated format, the mean satisfaction is higher for closed-ended tasks ($M = 3.616$, $SD = .218$) than for open-ended tasks ($M = 3.465$, $SD = .252$). On the animated format, the mean satisfaction is higher for open-ended tasks ($M = 3.729$, $SD = .255$) than closed-ended tasks ($M = 3.446$, $SD = .220$). Figure 24 below provides the mean satisfaction for closed-ended and open-ended tasks for each format. In Figure 25, we see that for closed ended tasks, the standard format has a slightly higher satisfaction over the other formats, next was the clustered, then non-animated and last the animated format, which subjects had the least satisfaction towards. In Figure 25, we see that for open-ended tasks, both standard and clustered have similar mean satisfaction, followed by animated, and then non-animated. Satisfaction was greater for animated visualization for open-ended tasks over closed-ended tasks.

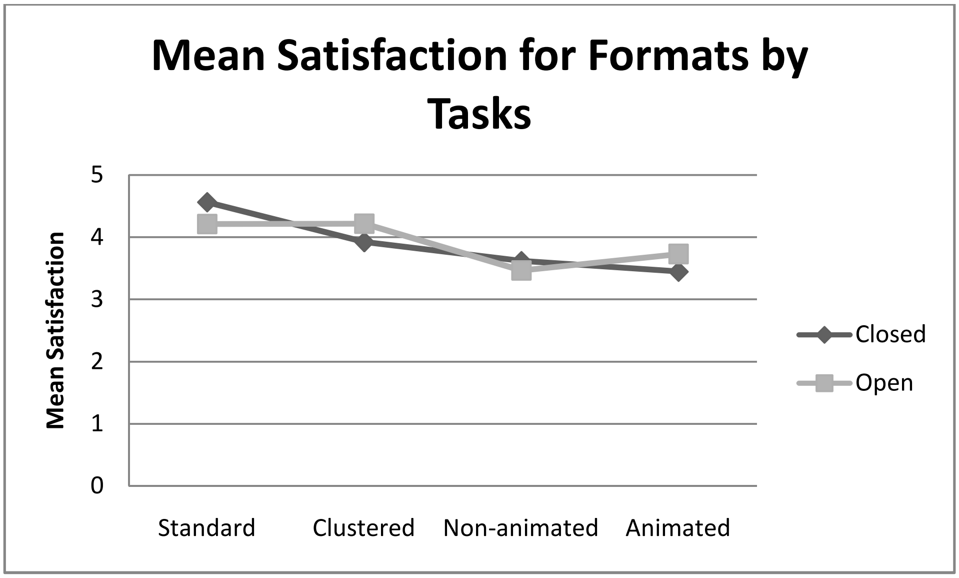


Figure 25. Estimated Marginal Mean Satisfaction for Formats by Tasks

8.6 Qualitative Analysis

Qualitative data was collected through protocol analysis to gain a deeper understanding of the performance and satisfaction outcomes of the study. We want to identify if the actual measures match with what participants went through when completing the experiments. Through the qualitative data, we are able to identify other issues that may not be determined through the quantitative performance measures. In addition, we are able to determine how subjects interpreted the different cues while using the prototype system. We can gain a better understanding of which cues were useful scents for subjects to help find relevant search results. Based on the subjects' verbalized feelings and actions, we were also able to determine the source of their satisfaction (or dissatisfaction) with the tool.

Five subjects in each format group were randomly selected to participate in the protocol analysis (Refer to Section 7.2 for the procedure for protocol analysis and instructions for subjects). Subjects were provided instructions on how to verbalize their thoughts and actions. Each subject was provided a headset. While performing the experiment, their voice and the actions they took were recorded using a software application called WebEx.

The WebEx videos were analyzed and coded by two independent experts. Based on protocol analysis, the videos were transcribed into segments by episodes or tasks (Ericsson and Simon, 1993). Each activity that a subject completed was identified. Therefore, a high level summary for each task that a subject completed was identified. After analyzing the videos and transcribed segments, codes were identified that were particular to this study.

As discussed in Section 7.2, protocol analysis data was collected to identify why subjects choose particular links/categories to help answer the tasks. The data was coded to remove irrelevant information that was not pertinent to this study. For example, it was not necessary to code how the subjects answered the demographic data questions. The process of coding involved mapping the data to the performance measures. We first viewed the videos and then identified cues to code the videos. The codes identified for this study are “cue”, “trial and error” and “satisfaction”.

The code of “cue” was selected to identify which cue (or scent) helped subjects pick their link/category on the given format. The code of “cue” was separated into different subsections based on the various cues that were provided in the format (keywords -title of category, animation, color, size, title of website, relevancy, random). We used random as a subsection of cue since subjects may not have used any cues to help pick the answer but a random selection. The code “trial and error” was selected to understand the amount of difficulty or effort that a subject went through to find the answer. The more “trial and error” a subject had in finding the answer would show that he/she had more difficulty or effort using that format to find the answer to the task. This could also be represented to show cognitive effort. The more “trial and error” a subject encountered, then the more cognitive effort that was expended. The last code selected for this study was “satisfaction”. Comments from subjects while using the tool provide evidence to positive or negative satisfaction toward the tool. Therefore, the “satisfaction” code was separated into two subsections, positive and negative.

The coding values used for this study are as follows:

- Cues (keyword or title of category, animation, color, size, title of website, relevancy, random)
- Trial and Error (how many websites were viewed prior to selecting the answer)
- Satisfaction (positive, negative)

Table 30 provides the results of the coded values from the qualitative analysis. Each column represents each coded value for this study. Each row represents each video of a subject. The column of “cue” provides the subsection (keywords (title of category), animation, color, size, title of website, relevancy, random) and number of times that cue was found in the video. The column of “Trial and Error” provides the number of “trial and error” that a subject encountered for each tasks (closed-ended and open-ended). The column of “satisfaction” provides the negative and positive comments that were found for each video. Table 31 provides a summary of overall results of the coded values for each format.

Table 30. Results of Coded Values for Qualitative Analysis

Format	Cues	Trial and Error	Satisfaction
Standard	Title of website - 14	Closed – 6 Open - 0	Used the “Find” tool to look for keywords on the results page Positive – 0 Negative - 0
	Title of website - 17	Open – 3 Closed - 6	Positive – 0 Negative - 0
	Title of website – 23	Open – 9 Closed - 7	Used the “find” tool to look for keywords on the results page Positive – 4 Negative - 0
	Title of website - 22	Open – 9 Closed - 7	Positive – 3 Negative - 0
	Title of website - 12	Open – 5 Closed - 3	Used the “find” tool Positive – 0 Negative - 0
Clustered	Random – 4 Title - 11	Open – 6 Closed - 7	Positive – 1 Negative - 5
	Title – 15 Relevancy - 1	Open – 3 Closed – 9	Negative – 5 frustrating at times – was not helpful to look for detailed information Positive - 0
	Title - 17 Relevancy - 3 Random - 2	Open – 9 Closed - 13	Positive – 1 Negative - 4
	Title - 16 Relevancy - 3	Open – 7 Closed - 12	Positive – 3 Negative – 3
	Title – 16 Relevancy – 4 Random - 4	Open – 6 Closed - 13	Positive – 1 Negative - 2
Non-animated	Title - 9 Title and Relevancy – 7 Random - 3	Open – 5 Closed - 15	Negative – 6 Circles overlap; can’t see the title clearly
	Title – 10 Title and Relevancy – 7	Open – 6 Closed 13	Negative – 4 Positive - 0
	Title - 12	Open – 4	Negative – 5

Format	Cues	Trial and Error	Satisfaction
	Title and Relevancy - 9	Closed - 12	Positive - 4
	Title – 7 Title and Relevancy - 7	Open – 4 Closed - 5	Negative – 3 Positive - 0
	Title – 8 Title and Relevancy – 4 Random - 6	Open – 6 Closed - 12	Negative – 3 Positive - 2
Animated	Title- 20	Open - 5 Closed - 18	Negative – 7 Positive- 0
	Title and blinking – 14 Title and blinking – 9 Title - 6	Open – 10 Closed - 21	Negative – 14 “I would have liked to write in my own queries” Would have liked to see the entire title, was not convenient, circles are so small to read the title, wish I could see the entire title Positive – 5 so many choices to look at; did give me a lot of choices, did provide a lot of information;
	Blinking - 8	Closed - 3 Open – 0	Picked the wrong query a few times before realizing that I need to pick the right query Positive – 2 Negative - 4
	Blinking – 15 Random – 2 Title - 7	Closed – 13 Open - 5	Spent 10 minutes looking at the wrong query results page – used 6 T/E and 4 Blinking, 3 random; 1 keyword
	Blinking – 14 Title - 5	Closed – 12 Open - 4	Positive – 2 Negative - 6

Table 31. Summary of Coded Values for Qualitative Analysis

Format	Cues	Trial/Error	Satisfaction
Standard	Title – 88	Closed - 26 Open - 29	Positive - 7 Negative - 0
Clustered	Title – 75 Relevancy – 11 Random - 10	Closed - 37 Open - 25	Positive - 6 Negative - 19
Non-animated	Title – 46 Title and Relevancy – 34 Random - 9	Closed - 57 Open - 25	Positive - 6 Negative - 21
Animated	Title – 38 Blinking – 37 Title and blinking –23 Random - 2	Closed – 67 Open - 24	Positive - 9 Negative - 31

8.7 Summary

In this chapter, we analyzed the data that was collected from the experiments conducted for this study. From the statistical analysis, we found support for H1a-H1b, where standard formats match closed-ended tasks. We found no evidence to support H1c-H1f, thus indicating visualization formats do not match open-ended tasks. We also conducted a multiple regression statistical analysis and correlations to determine the effects of cognitive effort on effectiveness, efficiency and satisfaction. We found support that cognitive effort of closed ended tasks had an effect on effectiveness and satisfaction; however cognitive effort for open-ended tasks had an effect on efficiency and satisfaction. We also coded the qualitative data into different code as it pertained to this study, which is discussed further in the next chapter.

CHAPTER 9

DISCUSSION AND SUMMARY

9.1 Closed-ended Tasks and Formats

For closed ended tasks, there was a significant difference between the formats. Subjects did better on the standard format (than the clustered format, non-animated and animated visualization. This provides support for H1a, H1b and H1c, where textual formats match closed-ended tasks. Subjects have to search for specific questions and standard formats provide results to closed-ended tasks towards the beginning of the results page; thus leading to a successful completion of the task. Therefore, the visualization formats was not suitable for closed-ended tasks since the first few relevant results may be scattered across different categories and become hard to find quickly. This supports the cognitive fit theory (CFT) which states that symbolic formats similar to standard formats will match symbolic tasks, which are specific similar to closed-ended tasks. According to information foraging theory, the title of the website is a significant scent that enabled subjects to find relevant results quickly.

Based on CFT, when there is a fit between format and task, positive outcomes will be obtained. Since there is a match or fit between closed-ended tasks and standard formats, cognitive effort for closed ended tasks is significant on effectiveness and satisfaction. When cognitive effort decreased, then effectiveness and satisfaction increased. This shows as subjects used less cognitive effort due to a match with format and task, their performance in effectiveness increased and their satisfaction with using the format increased. There was no significant difference in efficiency for closed-ended tasks. We believe that subjects were not concerned with time but rather wanted to ensure that the correct answers for the closed-ended tasks were found.

As subjects were completing closed ended tasks on visualization formats, their cognitive effort increased, their effectiveness decreased, and their satisfaction decreased. This provides support for CFT that subjects had to utilize more effort on visualization formats while completing closed-ended tasks since there is no match or fit between visualization formats and closed-ended tasks. Since closed-ended tasks were specific, it was easier to find answers on a standard textual format where results were listed in the first few results compared to a visualization format, where subjects have to find answers in different relevant categories.

9.2 Open-ended Tasks and Formats

For open ended tasks, subjects did better on the non-animated visualization than the standard format, clustered and visualization. Although subjects had a lower cognitive effort using non-animated visualization than the other formats, there was a slight difference in cognitive effort between the visualization formats and textual formats; however, the p-value was not significant to support the hypotheses. The qualitative discussions show that the prototype design for the visualization formats need to be improved and could possible lead to the slightly higher p-value that was obtained. Pairwise comparisons between the four formats also showed no significant difference.

Therefore, there was no support for H1d, H1e, and H1f, which states that visualization formats, specifically animated visualization match or fit with open-ended tasks

We claim that one of the reasons that subjects may have a lower cognitive effort using non-animated visualization format is that the relevancy number was associated to each category, which was not available on the animated visualization. In the animated visualization, only the top three relevant categories were animated. Since answering the open-ended tasks required subjects to browse through additional relevant categories to find the answers, subjects became confused as to what category was relevant after browsing the first three relevant (blinking) categories. The qualitative results found that although subjects felt the titles of categories overlapped with each other, in the non-animated visualization, they were able to browse through more than just three top relevant categories to find the answers. In the animated visualization, subjects were not able to recognize which category was relevant after browsing the top three relevant categories. A way to improve the design of the animated visualization is to include animation for relevancy on more categories than the top three relevant categories.

For open-ended tasks, cognitive effort had a significant difference on efficiency and satisfaction. When cognitive effort decreased, then satisfaction increased. When cognitive effort increases, then efficiency decreases. This supports the notion that when cognitive effort increases, the amount of time it takes to complete the tasks is longer than when there is a decrease in cognitive effort. This supports H3b.

There was no significant difference for cognitive effort on effectiveness for subjects completing open-ended tasks. The reason for no significant differences in effectiveness for open-ended tasks could be related to browsing for general topics can be subjective to the subject’s understanding of the question. There was no right or wrong answer but since relevant websites were already determined by three experts, there was no evaluation of what a person should focus on when answering these open-ended questions.

Although there were no significant differences on effectiveness for individuals completing the open-ended tasks on visualization formats, the qualitative data showed that subjects had slightly more “trial and error” (68) than subjects using the standard formats (57). Subjects in the non-animated visualization had less “trial and error” for open-ended tasks (25) than subjects using the animated visualization (42). This is a vast difference in completing open-ended task on non-animated versus animated visualization.

Based on information foraging theory, cues are scents that individuals use to help find relevant information quicker. Animation was a significant cue that we were testing in this study. In the animated visualization, the scent of animation was limited to the top three relevant categories only. Therefore, subjects had to utilize the only other scent, which were keywords (title) of the category available in the animated visualization after the scent of animation disappeared. The scent of keywords (titles) alone did not seem to be strong enough to allow for subjects to find results in open-ended tasks. Subjects in non-animated visualization format were able to choose results based on two scents, which were relevancy and the title of the category, since all the categories were labeled with the relevancy number. When answering open-ended tasks, subjects had to review more than three categories to find results, therefore, the subjects in the animated visualization did not know which category was relevant to the search after the scent of animation disappeared. The scent of animation or relevancy number seems to be very powerful cue.

If titles or keywords on categories are not provided with a scent of relevancy, then the strength of titles as a scent becomes weak in finding information faster. The design of the animated format should be improved to allow for more categories to be labeled with the scent of animation. In addition, there was less trial and error for open ended tasks across all formats. As discussed earlier, open-ended tasks were subjective. Subjects had to interpret the open-ended tasks and determine relevant websites to answer the questions based on their interpretation. Subjects in the standard formats picked more relevant websites due to the nature of picking the first few results on the results page.

9.3 Control Variables

The two control variables in the study that were statistically significant were cognitive style and experience. Cognitive style was controlled while testing the effects of format on cognitive effort required. The cognitive style measured an individuals' predisposition to either analytic or intuitive. It was stated that an individual with an analytic cognitive style would have a predisposition to standard formats and individuals with an intuitive cognitive style would have a predisposition to visual formats. However, results showed that intuitive individuals tended to have a lower cognitive effort required when completing both tasks on any format. This could be explained by the fact that intuitive individuals are more flexible and are able to use different formats to find results to the tasks. They are able to make judgments and conclusions based on a global view. Although this type of person was indicated to perceive a lower cognitive effort towards visualization formats, they could also perceive a lower cognitive effort required using the standard format since they have a predisposition towards variety, new ideas and openness to different experiences (See Table 6). They might have a more open-mind to find results using any format they are provided.

The control variable of experience on the Internet was statistically significant when testing for the effects of cognitive effort required on performance and satisfaction. When subjects had higher experience levels on the Internet, then performance and satisfaction increased. Subjects with higher levels of experience on the Internet would be able to navigate the formats and find accurate results to the tasks. They use the Internet more often than someone that is inexperience, so they might use a search engine often to search for information; thus enabling them to find the accurate results to the tasks. They might also been satisfied with using the format that they were randomly assigned since they are able to identify formats which help with navigating through vast number of results. Their performance might have increased since they could to find the results on the format easier than someone who is inexperienced on the Internet. None of the other control variables had any significance on the performance and satisfaction measures. The interaction of the control variables with the independent variable also did not produce significant results.

9.4 Design of the Tool

Subjects commented on the design of the prototypes and there was a mixed reaction to each tool. Although subjects did not have a good satisfaction rating for the visualization formats, there were comments such as "It was interesting but pretty difficult to navigate", "I think it is a good concept….but sill clunky….I could get used to with time", I thought that it was a good tool, but a little difficult to navigate pending on the topic of research" and "It would be a good tool if some things were changed. I could not read all of the writing around the circles (white was difficult and some of the words over lapped others). It would make more sense to me if the circle sizes matched with the relevancy, (the more relevant means the bigger the circle). And lastly, the drag button (blue up and down arrow bar) was also confusing."

Further comments regarding the visualization formats showed that the design of the tool could be improved. Based on information foraging theory, scents are very important to help individuals find information; however the strength of a scent enables individuals make decisions on whether to pursue a patch of information or to move on. It is possible to have various scents; however if the scent is not strong, then the design should be improved.

In this study, the title of categories was not an appropriate size to read. Therefore the scent of titles, which is a strong cue, was not being utilized to its fullest extent. Some of the categories overlapped with each other which made it confusing for subject. "The small circles were too hard to read and even when scrolling over them, the text would conflict with the other circles and was not readable. After clicking the circle, the whole process of it spinning around seemed like a waste of time and was annoying" and "long text gets jumbled together in the circle can't see it" and "The text was hard to see at some points".

Each cue on a format is important and can be a powerful scent for individuals to find information. Powerful scents, such as titles and keywords, should be appropriately identified and have a strong usability; otherwise the scent becomes weak and useless for its purpose which is to help individuals find information faster. In this study, the design of the format in regards to titles and layout of circles should be improved to ensure that the strength of the scent is achieved. In the animated visualization format designed for this study, the scent of animation should be improved to have a strong usability.

Subjects felt the clustered and visualization formats was difficult to understand. For clustered formats, “When looking for a simple answer, this engine proved to be too complex.” It seems as though it may be helpful at times but it is a little too complex as opposed to something like Google where the question can simply be typed in and you can usually get an answer relatively quickly.” For visualization formats, “I thought it was too complex.”, “It’s so complex and way too hard”, “… the interface was too complex.”

9.5 Appropriating the Tool Faithfully?

In the qualitative data analysis, WebEx videos showed that some subjects (3 out of 10) did not appropriate the tool faithfully to promote the technology’s spirit (DeSanctis and Poole, 1994). In other words, individuals were not using the tool as it was intended to be used by the designer. Subjects using the animated visualization and non-animated visualization did not always pick the correct queries to find the results for tasks. One subject (out of 5) using the animated visualization did not utilize the cue of animation to pick the initial categories. The subject picked categories based on title or keywords associated with each circle. This decreased their efficiency and may have led to their dissatisfaction and low performance scores.

9.6 Useful Cues for each Format

Subjects that did “appropriate the tool faithfully,” or use the tool correctly, for animated visualization utilized animation as a cue to select their initial categories. However, after selecting the first three categories, subjects did not know which category to pick based on relevancy. As discussed earlier, in the animated visualization, the rest of the categories were not labeled with the relevancy number as in the non-animated visualization. Therefore, subjects had to rely on the title or keywords of categories.

In the standard format, the qualitative analysis showed that subjects picked results based on whether the title of the website matched the task. Keywords and titles are important scents that should not be dismissed in designing information formats. Subjects in the visualization formats commented that the title of the websites was cut off which hindered them from seeing the entire title which is a significant cue in this study. “You weren't able to see the entire website title/article title which made it hard to make a decision as to chose it or not.” “Being able to school [scroll] through titles and read a little caption about the article is more helpful to me.” “It was hard to read some of the Titles at times.” Sometimes it may be hard to read the titles when they are written smaller and to recognize the significance of the circles being in size order. “I thought the words were too small and they were clustered together which made it even tough to read. Also, many of the titles couldn’t be seen because they were cut off. Other than this, I thought that it wasn’t too bad.” Therefore, titles/keywords are an important cue.

In the qualitative results, we saw that subjects in the clustered and non-animated visualization used the cue of keywords of the title to select categories in addition to the relevancy number. Subjects in the non-animated visualization did better on the open-ended tasks compared to all other formats because it seemed that a combination of relevancy number and title or keyword of the category helped subjects pick the correct websites to answer the tasks. In the standard format, the results are provided in a ranked list based on relevance of the results. Therefore, subjects in each of the groups (except the animated format group) had relevancy or rank provided for all results and categories. For the clustered and visualization formats, results are distributed across different categories and the use of relevance for each category would be useful. In the qualitative results, we noticed that subjects in the animated visualizations began to pick titles/keywords that matched the tasks they were working on after the scent of animation for the first three relevant categories disappeared. Since the performance of subjects in the animated formats was lower than all the other formats, the use of titles/keywords alone may not be sufficient for higher performance for animated visualization. The cues of titles/keywords, relevancy or animation are important cues and together they provide a powerful scent. However, they must be appropriately utilized and fully noticeable as described in Section 7.4. In the next section, we provide comments by subjects that describe the tools and cues used in this study.

9.7 Subject Comments regarding Formats

Subjects were given an opportunity to provide comments regarding the format they were using. The following table provides a small set of comments for each format.

Table 32. Comments from Subjects regarding Formats

Standard	• Did not particularly like it, a little irritating. • It turned up some rather irrelevant entries, could have been more organized and direct. • It was rather intuitive. • I don't know what the query was so I don't know if the results were ordered in some way but they did not appear to be, I went close to the end of the page for some answers. That was the only problem I had. Otherwise, it's a familiar interface so it was simple to view and use. • Actually I had limited ability to search. • I think the interface used here was great. I was not able to tell the search engine used to provide the results but could tell they were from different search engines and some of them gave better results than others. • I feel this could be used for finding general information quickly. The excerpts were helpful in identifying which page was the best fit. Detailed information was more difficult to pinpoint. • Finding information was a little hard because the provided website descriptions weren't very detailed. • Maybe 3-5 lines of actual text from the website would make searching easier, rather than 1-2 lines. • It was ok, but I like Google better. • Some of the searches were very easy but others were very hard. If I could have changed the text to search it would have been easier to find what I was looking for. • I liked the interface tool because it allowed me to search through all the information it could find based on the keywords. • I thought that the interface tool had good results but some of the findings were irrelevant. A lot of the good websites were at the top which was good. • I'd rather use Google.
Clustered	• I really didn't enjoy using it. It was extremely frustrating and made me want to end the questions quicker. It's way easier to type in exactly what you're looking for, rather than have to look through a list. Being able to school through titles and read a little caption about the article is more helpful to me. I didn't like the clusters and I didn't like having to click on each one. • The search engines are very helpful. Detailed information for a question took me a longer time to search • would not use it in the future • A bit time consuming

	- It was user friendly and not too complicated to understand for a first time user. Easy way to navigate multiple search query results as well as pages. - It was awkward to use and I felt that since I didn’t get a chance to put a specific description of what I wanted I got a lot of unnecessary URLS - I really didn't enjoy using it. It was extremely frustrating and made me want to end the questions quicker. It's way easier to type in exactly what you're looking for, rather than have to look through a list. Being able to school through titles and read a little caption about the article is more helpful to me. I didn't like the clusters and I didn't like having to click on each one. - Having various links for a certain topic was helpful! - It was an easy tool to use for basic questions. However, when the questions became more complex and more detailed information was needed it became difficult to find an appropriate source quickly and easily. - A bit confusing because everything is categorized, and sometimes I can't find the answer from a category with the highest relevancy. - I liked using the tree format for the first [open-ended tasks] section; it was helpful and provided me with very quick and organized results. For the second section it was extremely frustrating and it made me want to quit searching. - It was very helpful to collect large and complex data. Very useful for long research projects. - I think it is too complex to search anything. Say someone wanted to go to search something quickly and it was a general knowledge questions, how would they know exactly what query to use. It was okay for us during this experiment because it was given but other than that I think it would be too rigorous to get a task done
Non-animated	- I thought that it was a good tool, but a little difficult to navigate pending on the topic of research. - It was too complicated to figure out. It wasn't very user friendly. I still think that Google is much easier. - It will be a great tool for research papers. - It was very helpful for some topics yet it was too broad of a search a times and hard to find what was truly relevant - The small circles were too hard to read and even when scrolling over them, the text would conflict with the other circles and was not readable. After clicking the circle, the whole process of it spinning around seemed like a waste of time and was annoying. - I liked the beginning where there was a drop down box to pick a category. That circle confused me at first. I did not realize that

	there were ratings according to relevance for each category. It was easier to use after I got use to it. • It would be a good tool if some things were changed. I couldn’t read all of the writing around the circles (white was difficult and some of the words over lapped others). It would make more sense to me if the circle sizes matched with the relevancy, (the more relevant means the bigger the circle). And lastly, the drag button (blue up and down arrow & bar) was also confusing. • I think it is too complex to search anything. Say someone wanted to go to search something quickly and it was a general knowledge questions, how would they know exactly what query to use. It was okay for us during this experiment because it was given but other than that I think it would be too rigorous to get a task done. • I would not use this tool over google.com or yahoo.com because of the issue stated above. • I thought it was too complex. The text was hard to see at some points and there wasn’t information displayed when the search came up. • very cool • I was getting frustrated that it was taking me so long to answer the 1st set of questions [open-ended tasks] • long text gets jumbled together in the circle can’t see it • The interface is clunky-- too slow. Additionally, the response time when clicking on a link to display is slow or occasionally unresponsive. • It was frustrating not being able to type in my query. Also, the interface was too complex. • I would not use this tool for searches. Google and the library are easier • The tool needs to be more streamlined and provide greater detail on the sites in the folders. Some of the links were completely dead and lead to nothing, others had no material pertaining to the folder they were located in. The site consistently lagged and required far too much time to get from one subject matter to the next.
Animated	• the text was too small and hard to read with the varying colors behind the white text • I like the blinking circles, easy to see where to go • It should be bigger...I could barley read the text and the Whole headline was cut off. • the fonts should no overlap, it should be more clear and different colors should be used • The tool is pretty good • Too confusing. The color is not aesthetically pleasing at all.

	• It was difficult to find some of the answers to the questions. • It was annoying, it took too long to load the pages, the keywords on each circle were vague and gave me no clue as to what types of results I would find there. Also, it was unnecessary to list all the results in the large circle; most of them went unused in my case • Very useful but too much information, thus a little bit confusing! • It was a little confusing at first because what appeared to be the source where the answer was, I couldn't find it. What appeared to be the 2nd most likely place to where the answer was, I found easy. That was only for 1 question though but still. • I think that the tool is very different from any other search engine that I have ever used before, and it takes a little time to get used to how it works. Sometimes it may be hard to read the titles when they are written smaller and to recognize the significance of the circles being in size order. It looks like it's part of the design less based on the relevance of the information. I think it is admirable to create such an innovated tool, but with any variation or upgrade it will take time for people to feel comfortable with the tool. • The tool is pretty good • Most relevant content did not appear in the most relevant bubble • You weren't able to see the entire website title/article title which made it hard to make a decision as to choose it or not. • Very difficult to use. Had a hard time finding the specific answers because it was difficult to sort through the results. • Very useful • I think that it's a good concept that provides for a higher level of interactivity than most engines. However it is still clunky probably owing to it still being in development. It might also be difficult for older people who are not comfortable with computers other than that I think it is something I could get used to with time. • I found it relatively easy. It helped me answer the questions effectively. Although it took longer then I normally would to use a search enjoyed, I would definitely use it again. • I felt like I could have found the answers much faster by using a different search engine • It was interesting but pretty difficult to navigate. The query took too long as well. • I thought that it was way too difficult and took too much time to answer the questions • It was useful at times

Overall, subjects felt that the standard format is similar to Google. Therefore, the standard format prototype was appropriately designed to identify a baseline for this study. Subjects in this group were used to using the tool and were able to find answers quickly especially for closed-ended tasks. For clustered formats, subjects had a mixed reaction that ranged from helpful ("very helpful", "user friendly and not too complicated", each way to navigate multiple search query results", "various links for a certain topic was helpful" to complicated ("would not use it in the future, "a bit time consuming", "extremely frustrating". For the visualization formats, there were more comments regarding the design "It would be a good tool if some things were changed", which seemed to hinder a positive response for the tool. Many of the comments were based on the complication of the tool, which can be attributed to the fact that many of the subjects use Google as their main search tool. "I would never use this tool again", "too complex", "getting frustrated finding results". However, there were also positive comments: "it will be a great tool for research papers", "helpful for some topics yet it was too broad for a search at times". For animated visualization, the comments were similar based on the design of the tool: "It should be bigger...I could barley read the text and the whole headline were cut off" and "the text was too small and hard to read with the varying colors behind the white text". However, some subjects were happy with the scent of animation: "I liked the blinking circles, easy to know where to go", "the tool is pretty good" and "very useful".

9.8 Summary

In this chapter we discussed the findings for this study. Cues can be powerful scents if they are strong and useful. Animation is a powerful scent; however in the animated visualization format designed for this study, the scent of animation could have been more fully utilized; therefore, this potentially powerful cue was not sufficiently strong to result in significant performance improvements. Since subjects had to browse for information past the first three animated categories, the scent of animation did not help users to identify other relevant categories. Subjects also had mixed reactions to the tool; however it was evident that in future studies the design of the tool should be improved to allow subjects to easily read titles and keywords, which were also identified as powerful cues.

One limitation of this study is that the four prototypes differ with regard to information content contained within the available cues. This may have influenced the results. Future studies should convey the same information across the four formats. In this study, the standard format utilized the cues of website titles and rank in the ordered list; however, the rank was not labeled next to each result, which should be incorporated in future prototypes. The clustered, non-animated and animated visualization formats utilized keywords of categories, relevancy or animation with various cues, such as color and size. In future prototypes, the color utilized in each of the formats should be similar. In addition, there should only be one size for all circles to represent the categories in both the non-animated and animation visualization formats. In this way, we would be able to reduce the effects of any other potential cues that may have dampened the effects of animation as a strong scent.

Future studies should also compare the clustered format group versus animated visualization. In both of these formats, the results are categorized into clusters; however the display and cue for relevancy is the only difference between the two formats. In this way, we can identify whether a visualization map or animation can help subjects find results for particular tasks.

In addition to improving the design of the tools, other factors should be considered for future studies. The qualitative results showed that subjects (3 out of 5) in the standard format group browsed results using the “find” feature on Internet Explorer by entering keywords that matched the task. Since all three hundred results were on one page, this feature may have helped subjects in this group filter through the results rather quickly. In future studies, the “find” feature should be disabled to ensure that subjects using a standard textual format do not have an additional feature to help find results quickly.

Future studies should also consider how to appropriately score open-ended tasks. Open-ended tasks are subjective and sometimes there is no right or wrong answers. For the purposes of this study, we took the approach of measuring effectiveness or the score for open-ended tasks based on the relevant results selected by experts.

Another factor to consider for future studies is the measure for cognitive effort required. We tested our theory based on the theoretical model and measured cognitive effort required based on a subjective survey for cognitive decision and cognitive convenience. Future studies should try to measure cognitive effort based on an objective measure such as the number of clicks that a subject performs per task.

Based on information foraging theory, this study confirmed that scents are powerful cues; however each scent is only as powerful as the strength and utilization of each cue. The cues on the formats used in this study should be improved to ensure the full capacity of powerful scents. In next chapter, we discuss the contributions that this study has made.

CHAPTER 10

CONCLUSION

This study makes several contributions. First, we developed a visualization interface for search engine results using cues such as color, size, and animation. Since color and size were previously used in visualization, we studied the effects of animation on visualization of search results. Second, we provided insight into animation as a useful cue or scent for browsing relevant information. Although we did not find significant evidence of animation as a useful cue in this study, future studies should try to focus on using animation on more relevant categories than just three as seen in this study. According to the qualitative data, the effects are animation is useful as subjects did utilize the cue to find relevant categories; however the number of categories with animation was not sufficient. Third, we identified that closed-ended tasks fit with a standard textual format and that open-tasks do not fit with visualization formats. However, for open-ended tasks, subjects had a lower cognitive effort using the non-animated visualization format than the others; however the mean cognitive effort for both visualization formats (animated and non-animated) did not show significant findings over the combined textual formats. This leads to the last contribution of this study. The results of this study should help web designers understand the effects visualization and use of different cues. As more and more information is available to individuals, finding relevant information in a timely manner becomes very crucial; however, if the formats that provide information to individuals are not useful then it defeats the purpose.

Chen (2005) found that one of the unsolved problems of information visualization is aesthetics or identification of features that are insightful and visually appealing rather than utilizing features that are just appealing to the user. According to information foraging theory, scents are powerful cues that can help individuals find information faster. The stronger the scent is in finding information, the more useful the scent will be for individuals. The results of this study should help designers of web search engines to understand the impact of designing visualization formats and using strong scents such as animation, keywords, relevancy or other features to enable faster and better searching performance. Designers may be able to use animation as a visual scent to reduce information overload; however, designers should be aware that the scent of animation should not disappear after the top three relevancy categories as designed in this study. Other scents such as keywords and relevancy numbers seem to be strong scents on formats to find information quicker. The scents in the visualization formats should provide sufficient meaning throughout the format, similar to how keywords (titles), relevancy and animation provide. The design of the formats should be simple to use and provide users with the ability to understand the entire format. As indicated in comments by subjects, the more the scents became weak or confusing to read, the less the scents became powerful and useful.

> *I like the blinking circles, easy to see where to go.* (Blinking circles refer to animation.)

> *It would be a good tool if some things were changed. I couldn't read all of the writing around the circles (white was difficult and some of the words over lapped others).*

It would make more sense to me if the circle sizes matched with the relevancy, (the more relevant means the bigger the circle). And lastly, the drag button (blue up and down arrow & bar) was also confusing.

I think that the tool is very different from any other search engine that I have ever used before, and it takes a little time to get used to how it works. Sometimes it may be hard to read the titles when they are written smaller and to recognize the significance of the circles being in size order. It looks like it's part of the design less based on the relevance of the information. I think it is admirable to create such an innovated tool, but with any variation or upgrade it will take time for people to feel comfortable with the tool.

To increase the growth of information visualization, Chen (2005) states evaluating the usefulness of visualization components through elementary cognitive tasks, such as recognition of cluster based on their proximity or identification of a trend. This study used browsing and closed-ended tasks to examine the usefulness of animation as a component of visualization information. Future studies can use cognitive tasks such as identifying an individual's ability to discover new connections or patterns between results on the prototypes developed in this study. Future studies should also try to utilize animation as a powerful scent in a further capacity for information visualization formats to utilize the effects of animation as a cue to help individuals find information faster.

REFERENCES

Abdel-Halim, A.A. (1980). Effects of higher order need strength on the job performance-job satisfaction relationship. *Personnel Psychology*, *33*(2), 335-347.

Abdi, H. (2007). Zscores. In Neil Salkind (Ed.) *Encyclopedia of Measurement and Statistics*. Thousand Oaks (CA): Sage.

Ackoff, R. L. (1967). Management misinformation systems. *Management Science, 14,* 147-156.

Agarwal, R., Sinha, A. P., & Tanniru, M. (1996). Cognitive fit in requirements modeling: A study of object and process methodologies. *Journal of Management Information Systems*, *13*(2), 137-162.

Allan, J., Leuski, A., Swan, R., & Byrd, D. (2001). Evaluating combination of ranked lists and visualizations of inter-document similarity. *Information Processing & Management, 37*, 435-458.

Allinson C.W. and Hayes, J. (1996). The Cognitive Style Index: a measure of intuition-analysis for organizational research. *Journal of Management Studies*, 33, 119-135.

Ames, C., & Archer, J. (1988). Achievement goals in classroom: Students' learning strategies and motivation process. *Journal of Educational Psychology, 80*, 260-267.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Bawden, D. (2001). *Information Overload.* London: LITC Publications.

Benbasat, I., & Dexter, A. S. (1986). An investigation of the effectiveness of color and graphical Information presentation under varying time constraints. *MIS quarterly, 10*(1), 59-82.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (Eds.). (1999). *Readings in Information Visualization: Using Vision to Think.* San Francisco, CA: Morgan Kauffmann Publishers.

Card, S. K., Moran, T. P., & Newell, A. (Eds.). (1983). *The Psychology of Human-Computer Interaction.* Hillsdale, NJ: Lawrence Erlbaum and Associates.

Card, S. K., Pirolli, P., Wege, M. V. D., Morrison, J. B., Reeder, R. W., Schraedley, P. K., & Boshart, J. (2001). Information scent as a driver of web behavior graphs: Results of a protocol analysis method for web usability. *ACM Transactions on Computer-Human Interaction, 3*(1), 498-505.

Chan, A. H. S., & Lee, P. S. K. (2005). Effect of display factors on Chinese reading times, comprehension scores and preferences. *Behavior & Information Technology, 24*(2), 81-91.

ChanLin, L. (2000). Attributes of animation for learning scientific knowledge. *Journal of Instructional Psychology, 27*, 228-238.

Chen, C. (2005). Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications, 25*(4), 12-16.

Chen, M., Hearst, M., Hong, J., & Lin, J. (October 1999). *CHA-CHA: A System for Organizing Intranet Search Results.* Boulder, CO.

Chi, E. H., Pirolli, P., & Pitkow, J. (2000, April 01-06). The Scent of a Site: A System for Analyzing and Predicting Information Scent, Usage, and Usability of a Web Site. Proceedings of the SIGCHI conference on Human factors in computing systems, Hague, Netherlands.

Chi, E. H., Pirolle, P., Chen, K., & Pitkow, J. (2001). Using Information Scent to model user information needs and actions and the Web. Paper presented at the Conference on Human Factors in Computing Systems, Seattle, WA.

Chung, W., Chen, H., & Nunamaker, J. F. (2005). A Visual Framework for Knowledge Discovery on the Web: An Empirical Study of Business Intelligence Exploration. *Journal of Management Information Systems, 21*(4), 57-84.

Coe, M. (1996) *Human Factors for Technical Communicators*. Canada: John Wiley & Sons, Inc.

Cooper-Martin, E. (1993). Effects of information format and similarity among alternative on consumer choice processes. *Journal of the Academy of Marketing Sciences, 21*, 239-246.

Cove, & Walsh. (1998). Online text retrieval via browsing. *Information Processing & Management, 24*(1), 31-37.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94*(2), 428-434.

Crapo, A. W., Waisel, L. B., Wallace, W. A., & Willemain, T. R. (2000). Visualization and The Process of Modeling: A Cognitive-theoretic View. Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining, Boston, MA.

Christen, M., Iyer, G., & Soberman, D. (2006). Job satisfaction, job performance and effort: a reexamination using agency theory. Journal of Marketing, 70(1), 137-150.

Cools, E., Van Den Broeck, H., & Bouckenooghe, D. (2006) The Cognitive Style Indicator: Development and validation of a new measurement instrument, Working Paper.

Cutting, D. R., Karger, D. R., Pedersen, J. O., & Tukey, J. W. (1992). Scatter/Gather: A cluster-based approach to browsing large document collections. Proceedings of the Conference on Research and Development in Information Retrieval, Copenhagen, Denmark.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality, 62*(1), 119-142.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268.

DeLone, W.H. & McLean, E.R. (1992). Information systems success: The quest for the dependent variable. Information Systems Research, 3(1), 60 – 95.

Dennis, A. R., & Carte, T. A. (1998). Using geographical information systems for decision making: Extending cognitive fit theory to map-based presentations. *Information Systems Research, 9*(2), 194-203.

Di Stephano, J. (2003) How much Power is Enough? Against the Development of an Arbitrary Convention for Statistical Power Calculations. *Functional Ecology*, 17(5), 707-709.

Dumais, A. R., Cutrell, E., & Chen, H. (2001). Optimizing search by showing results in context. Proceedings of the SIGCHI conference on Human factors in computing systems, Seattle, WA.

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society, 20*, 325-344.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Gagne, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, *26*, 331-362.

Gattis, L. (November 2002). Planning and Information Foraging Theories and Their Value to the Novice Technical Communicator. *ACM Journal of Computer Documentation*, *26*(4), 168-175.

Gill, T. G. (1996). Expert Systems Usage: Task Change and Intrinsic Motivation. *MIS Quarterly*, *20*(3), 301-329.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion, 24*, 175-213.

Gulli, A. & Signorini, A. (2005) The indexable web is more than 11.5 billion pages. Proceedings of the International World Wide Web Conference: Special interest tracks and posters of the 14th international conference on World Wide Web. Chiba, Japan.

Healey, C. G., Booth, K. S., & Enns, J. T. (1995). Visualizing real-time multivariate data using preattentive processing. *ACM Transactions on Modeling and Computer Simulation, 5*(3), 190-221.

Hearst, M. (1995). TileBars: Visualization of term distribution information in full text information access. Proceedings of the Conference on Human Factors in Computing Systems, Denver, CO.

Heo, N., & Sundar, S. S. (2001). Memory for Web Advertisements: Exploring Effects of Animation, Position, and Product Involvement. Paper presented at the International Communication Association Conference, Washington, DC.

Heo, N., & Sundar, S.S. (2000) Emotional Responses to Web Advertising: The Effects of Animation, Position, and Product Involvement on Physiological Arousal. Paper presented at the Advertising Division at the Annual Conference of the Association for Education in Journalism and Mass Communication, Phoenix, AZ, August 9-12.

Ho, J. & Tang, K. (2001) Towards an optimal resolution to information overload: An infomediary approach. Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work, Boulder, CO: ACM Press, 91-96.

Hoadley, E. (1990). Investigating the effects of color. *Communication of the ACM, 33*(2), 120-125.

Hong, E., & O'Neil, H. F. J. (1992). Instructional Strategies to Help Learners Build Relevant Mental Models in Inferential Statistics. *Journal of Educational Psychology, 84*(2), 150-159.

Hong, W., Thong, J. Y. L., & Tam, K. Y. (2003). Does Animation Attract Online Users' Attention? The Effects of Flash on Information Search Performance and Perceptions. *Information Systems Research, 15*(1), 60-86.

Hong, W., Thong, J. Y. L., & Tam, K. Y. (2004). The Effects of Information Format and Shopping Task on Consumers' Online Shopping Behavior: A Cognitive Fit Perspective. *Journal of Management Information Systems, 21*(3), 149-184.

Hunt, R.G., Krzystofiak, F.J., Meindl, J.R., & Yoursry, A.M. (1989)Cognitive Style and Decision Making. *Organizational Behavior and Human Decision Processes*, 44, 436-453.

Keller, K. L., & Staelin, R. (1987). Effects of Quality and Quantity of Information on Decision Effectiveness. *Journal of Consumer Research, 14*, 200-213.

Kim, K., Yoo, C. Y., & Stout, P. (2003). Has Animation Been Overused in Online Advertising? Effects of Animation and the Role of Affective Responses on Viewers' Perceptions of Banner Ads. Paper presented at The American Academy of Advertising 2003 Conference, Denver, CO.

Kosslyn, S. M. (1994). *Elements of Graph Design*. New York: W.H. Freeman.

Kuhlthau, C. C. (1991). Inside the Search Process: Information Seeking from the User's Perspective. *Journal of the American Society for Information Science, 42*(5), 361-371.

Lang, A., Borse, J., Wise, K., & David, P. (2002). Captured by the World Wide Web: Orienting to Structural and Content Features of Computer-Presented Information. *Communication Research, 29*(3), 215-245.

Lewandowski, D. (2005). Web searching, search engines and information retrieval. *Information Services & Use*, *25*, 137-147.

Li, H., & Burkovac, J. L. (1999). Cognitive Impact of Banner Ad Characteristics: An Experimental Study. *Journalism & Mass Communication Quarterly, 76*(2), 341-353.

Lin, X. (1997). Map Displays for Information Retrieval. *Journal of the American Society for Information Science*, *48*(1), 40-54.

Lyman, P. & Varian, H.R. (2000). How much information? School of Information Management and Systems, University of California at Berkeley, http://www2.sims.berkeley.edu/research/projects/how-much-info/internet.html

Mackinlay, J. D., Robertson, G. G., & DeLine, R. (1994). Developing Calendar Visualizers for the Information Visualizer. Proceedings of the 7th annual ACM symposium on User interface software and technology, Marina del Rey, CA.

Mao, J. & Benbasat, I. (2000) The Use of Explanations in Knowledge-Based Systems: Cognitive Perspectives and a Process-tracking Analysis. *Journal of Management Information Systems*, 17(2), 153-179.

Marchionini, G., & Shneiderman, B. (1988). Finding facts vs. browsing knowledge in hypertext systems. *IEEE Computer, 2*(1), 70-80.

Mayer, R. E., & Chandler, P. (2001). When Learning is Just a Click Away: Does Simple User Interaction Foster Deeper Understanding of Multimedia Messages. *Journal of Educational Psychology, 93*, 390-397.

McDonald, D. M., & Chen, H. (2006). Summary in Context: Searching Versus Browsing. *ACM Transactions on Information Systems*, *24*(1), 111-141.

McGraw, K. O., & McCullers, J. C. (1979). Evidence of a Detrimental Effect of Extrinsic Incentives on Breaking a Mental Set. *Journal of Experimental Social Psychology, 15*, 285-294.

Meyer, J. P., Becker, T. E., & Vandenberghe, C. (2004). Employee Commitment and Motivation: A Conceptual Analysis and Integrative Model. *Journal of Applied Psychology*, *89*(6), 991-1007.

Montazemi, A. R., & Wang, S. (1988). The effects of modes of information presentation on decision-making: a review and meta-analysis. *Journal of Management Information Systems*, *5*(3), 102-126.

Murray, B. & Moore, A. (2000). Sizing the Internet. Cyveillance. *www.cyveillance.com/web/downloads/Sizing_the_Internet.pdf*.

Nakakoji, K., Takashima, A., & Yamamoto, Y. (2001). Cognitive Effects of Animated Visualization in Explanatory Visual Data Analysis. Proceedings of the Fifth International Conference on Information Visualization.

Nickerson, R., Perkins, D. & Smith, E. (1985). *The Teaching of Thinking*. Erlbaum, Hillsdale, NJ.

Nielsen, J. (2000). *Designing Web Usability*. Thousand Oaks, CA: New Riders Publishing.

Olston, C., & Chi, E. H. (2003). ScentTrails: integrating browsing and searching on the Web. *ACM Transactions on Computer-Human Interaction, 10*(3), 177-197.

Owen, S. Budgen, D. & Brereton, P. (2006). Protocol Analysis: A Neglected Practice. *Communications of the ACM*, 49(2), 117-122.

Percy, L., & Rossiter, J. R. (1983). Effects of picture size and color on brand attitude responses in print advertising. *Advances in Consumer Research, 10*, 17-20.

Pereira, R.E. (2000). Optimizing human-computer interaction for the electronic commerce environment. *Journal of Electronic Commerce Research*, *1*(1), 23- 44.

Phillips, B., & Lee, W. (2005). Interactive animation: exploring Spokes-characters on the Internet. *Journal of Current Issues and Research in Advertising, 27*(1), 1-17.

Pinker, S. (Ed.). (1984). *Visual Cognition.* Amsterdam, The Netherlands: Elsevier Science Publishers BV.

Pirolli, P. (1997, March 22-27). Computational Models of Information Scent-Following in a Very Large Browsable Text Collection. Proceedings of the SIGCHI conference on Human factors in computing systems, Atlanta, GA.

Pirolli, P. (2003). Exploring and Finding Information. In J. M. Carroll (Ed.), *HCI Models, Theories and Frameworks: Toward a Multidisciplinary Science,* (157-191). San Francisco, CA: Morgan Kauffmann Publishers.

Pirolli, P., & Card, S. K. (1998). Information foraging Models of Browsers for very Large Document Spaces. Proceedings of the working conference on Advanced visual interfaces, L'Aquila, Italy.

Pirolli, P., Card, S. K., & Van Der Wege, M. M. (2001). Visual information foraging in a focus + context visualization. *ACM Transactions on Computer-Human Interaction, 3*(1), 506-513.

Pirolli, P., Card, S. K., & Van Der Wege, M. M. (March 2003). The effects of information scent on visual search in the hyperbolic tree browser. *ACM Transactions on Computer-Human Interaction, 10*(1), 20-53.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879-891.

Proctor, R. & Van Zandt, T. (1994). Human Factors in Simple and Complex Systems: Needham heights, MA: Allyn & Bacon.

Rieber, L. P. (1996). Animation as feedback in a computer-based simulation: representation matters. *Educational Technology Research & Development*, *44*(1), 5-22.

Roussinov, D., & Chen, H. (2001). Information navigation on the web by clustering and summarizing query results. *Information Processing & Management, 37*(6), 789-816.

Russo, J. E., & Dosher, B. A. (1983). Strategies for multi-criterion binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 676-696.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68-78.

Schick, A. G., Gorden, L. A., & Haka, S. (1990). Information overload: A temporal approach. *Accounting Organizations and Society, 15*, 199-220.

Schneider, S. C. (1987). Information overload: Causes and consequences. *Human Systems Management*, *7*, 143-153.

Schultze, U., & Vandenbosch, B. (1998). Information overload in a groupware environment: Now you see it, now you don't. *Journal of Organizational Computing and Electronic Commerce*, *8*(2), 127-148.

Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations. Proceedings of the 1996 IEEE Symposium on Visual Languages, 336 – 343.

Sinha, A. P., & Vessey, I. (1992). Cognitive fit: an empirical study of recursion and iteration. *IEEE Transactions on Software Engineering*, *18*(5), 368-379.

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119* (1), 3 –22.

Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences*, *34*(4), 771-797.

Spence, R. (2001). *Information Visualization*. New York: ACM Press.

Stasko, J., Catrambone, R., Guzdial, M., & McDonald, K. (2000). *An Evaluation of Space-Filling Information Visualizations for Depicting Hierarchical Structures.* Atlanta, GA: Georgia Institute of Technology, Graphics, Visualization and Usability Center.

Sundar, S. S., & Kalyanaraman, S. (2004). Arousal, memory, and impression-formation effects of animation speed in web advertising. *Journal of Advertising, 33*(1), 7-17.

Swash, G. (1998). UK business on the Internet. *New Library World, 99*(1144), 238-242.

Thomas, B. H., & Calder, P. (2001). Applying cartoon animation techniques to graphical user interfaces. *ACM Transactions of Computer Human Interaction, 8*(3), 198-222.

Tufte, E. R. (2001). *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

Turetken, O., & Sharda, R. (2004). Development of a fisheye-based information search processing aid (FISPA) for managing information overload in the web environment. *Decision Support Systems, 37*, 15-34.

Turetken, O., & Sharda, R. (2005). Clustering-based visual interfaces for presentation of web search results: an empirical investigation. *Information Systems Frontiers, 7*(3), 273-297.

Turetken, O., & Sharda, R. (2007), "Visualization of Web Spaces: State of The Art and Future Directions" in *DATA BASE for Advances in Information Systems* Vol. 38(3), pp. 51-81.

Vallerand, R. J. (2000). Deci and ryan's self-determination theory: a view from the hierarchical model of intrinsic and extrinsic motivation. *Psychological Inquiry, 11*(4), 312-319.

Vessey, I. (1991). Cognitive fit: theory-based analyses of the graphs versus tables literature. *Decision Sciences, 22*(1), 219-241.

Vessey, I., & Galletta, D. (1991). Cognitive fit: An empirical study of information acquisition. *Information Systems Research, 2*(1), 63-84.

Walsh, H.E., Kidd, M.G., Moum, T. & Friesen, V.L. (1999) Polytomies and the Power of Phylogenetic Inference. *Evolution*, 53(3), 932-937.

Wise, J. A., Thomas, J. J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., & Crow, V. (1995). Visualizing the non-visual: Spatial analysis and interaction with information from text documents. *IEEE Computer Society*, 51-58.

Witkin, H.A., Moore, C.A., Goodenough, D.R. & Cox, P.W. (1997). Field-Dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research*, 47, 1-64.

Yoo, C. Y., Kim, K., & Stout, P. (2003). Effectiveness of Animated Banner Advertising: Heirarchy of Effects Model. Paper presented at the American Academy of Advertising 2003 Conference, Denver, CO.

Yoo, H., Jang, D., Seo, K., & Lee, M. (2004). Retrieving images by comparing homogeneous color and texture objects in the image. *International Journal of Pattern Recognition, 18*(6), 1093-1110.

Zamir, O., & Etzioni, O. (1999). Grouper: A Dynamic Clustering Interface to Web Search Results. www.cs.washington.edu/research/clustering

Zhang, P. (2000). The effects of animation on information seeking performance on the world wide web: securing attention or interfering with primary tasks? *Journal of the Association for Information Systems, 1*(1), 1-28.

"Between-Subject Versus Repeated-Measures Designs" http://www.med.monash.edu.au/spppm/research/rda/Between%20subject%20vs%20repeated%20measures.htm retrieved on November 22, 2005.

"Structural Equation Modeling" http://www2.chass.ncsu.edu/garson/pa765/structur.htm retrieved on January 30, 2007.

Internet Activities (June 30, 2004). Retrieved December 05, 2005, from the World Wide Web: http://www.pewinternet.org/trends/Internet_Activities_4.23.04.htm

Number of Websites (1990-2008): http://royal.pingdom.com/?p=273 retrieved on 9/22/2008

APPENDICES

APPENDIX A: Sample Tasks

Sample Task

Question: What are the IT services that integrate business processes offered by the company NCS?

Closed-ended tasks

Question 1: What is the title of the article written by Ronen Feldman on text mining approaches?

Question 2: Which University is the 2008 Top National University in the US as indicated by USNews?

Question 3: Name the new video series that was created by the Official Visitor Site Greater Philadelphia organization to explore 24 of Philadelphia's finest neighborhoods.

Question 4: In what year was Temple University established?

Question 5: Who was the 32nd US President?

Open-ended tasks

Question 6: Describe the similarities and differences in the political views in education of Hilary Clinton, Barack Obama, and John McCain.

Question 7: Imagine you are writing a paper on music. Describe how the genre of "hip hop" in 2008 has influenced our culture.

APPENDIX B: Satisfaction Measures

Taken from Turetken and Sharda (2005) where they adapted a multi-item scale from Stasko et al. (2000)

Satisfaction Survey (Scale 1 = strongly disagree; 7 = strongly agree)

	1	2	3	4	5	6	7
Strongly disagree							Strongly agree

1. There are definitely times that I would like to use this system.
2. I would like this system available for my use all the time.
3. I found this system useful.
4. I found this system confusing to use.
5. I liked this system.

APPENDIX C: Cognitive Effort Measures

Obtained from Hong et al. (2004)

Cognitive Decision Effort
Please indicate the degree to which you would agree with the following statements by choosing a number from 1 to 10, where 1 indicates "strongly disagree" and 10 indicates "strongly agree."

1	2	3	4	5	6	7	8	9	10
Strongly disagree									Strongly agree

1. To complete the task, using this web site was very frustrating.
2. To complete the task, using this web site took too much time.
3. To complete the task, using this web site required too much effort.
4. To complete the task, using this web site was too complex.
5. I easily found the information I was looking for. *(R)
6. To complete the task, using this web site was easy. (R)

Cognitive Convenience
Please indicate the degree to which you would agree with the following statements by choosing a number from 1 to 10, where 1 indicates "strongly disagree" and 10 indicates "strongly agree."

1	2	3	4	5	6	7	8	9	10
Strongly disagree									Strongly agree

1. It is easy for me to move to the target page.
2. It is convenient for me to look for detailed information.

* This item was dropped due to the result of factor analysis.
R The items were reverse scaled.

APPENDIX D: Experimental Procedure

Experimental Procedure

-	Discuss and Sign Consent Form	3 min
-	Demographics survey	3 min
-	Cognitive Style Index questionnaire	7 min
-	Training	10 min
-	Task 1	15 min
-	Satisfaction questionnaire	3 min
-	Cognitive effort questionnaire	3 min
-	Task 2	30 min
-	Satisfaction questionnaire	2 min
-	Cognitive effort questionnaire	2 min

Total ~ 78 min

APPENDIX E: Final Survey

Subject Number ______________________________

Group Number ______________________________

1) What is your age?	15 - 20 ☐	20-25 ☐	25-30 ☐	30-35 ☐	35-40 ☐	>40 ☐

2) What is your major? ______________________________

3) What is your gender? ☐ Female ☐ Male

4) Is English your primary language? ☐ Yes ☐ No

5) How much experience do you have using search engines on the Internet?	Little (1-3 years) ☐	Moderate (3-5 years) ☐	Extensive (5 and above) ☐

6) How many times a week do you use search engines on the Internet?	Less than 1 ☐	1 to 10 ☐	10 to 20 ☐	20 to 30 ☐	More than 30 ☐

7) How much experience do you have writing queries on search engines?	Little ☐	Moderate ☐	Extensive ☐

8) Which search engine do you primarily use for finding information?
(Choose as many as applicable)

☐ Google ☐ KartOO
☐ AltaVista ☐ MSN
☐ Yahoo! ☐ Other ___________

Imagine you are working on a project or activity for work or school, Please indicate the degree to which you would agree with the following statements of your motivation to complete the activity by choosing a number from 1 to 7, where 1 indicates “strongly disagree” and 7 indicates “strongly agree.”

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
Trying hard is important to me.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
Getting the correct answers is the main goal of an activity.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
It is important that I do well in an activity.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
I want to succeed in an activity.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
I will perform my best regardless of how difficult the task.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
I always want to do my best in any activity.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7

Instructions:

Thank you for participating in this study.

You will be asked a series of questions. The results to a query for each question are provided on a search results page. Please read the instructions on viewing the results on the search results pages as indicated on the web page assigned to you based on your Group Number.

Write your answers in the space provided below each question. It is more important that you write the URL(s) from where you obtained the answers for each question.
Please record time as indicated on this sheet.
If you have any questions, please raise your hand and someone will assist you.

If you close the results page, please go to the Favorites menu on your Internet browser and select your group number that was assigned to you.

Time started: __________

Question 1: What is the title of the article written by Ronen Feldman on text mining approaches?

Answer

URL(s)

Question 2: Which University is the 2008 Top National University in the US as indicated by USNews?

Answer

URL(s)

Question 3: Name the new video series that was created by the Official Visitor Site Greater Philadelphia organization to explore 24 of Philadelphia's finest neighborhoods.

Answer

URL(s)

Question 4: In what year was Temple University established?

Answer

URL(s)

Question 5: Who was the 32nd US President?

Answer

URL(s)

Time completed: __________

Please indicate the degree to which you would agree with the following statements on the given information format that you used to complete the above tasks by choosing a number from 1 to 7, where 1 indicates “strongly disagree” and 7 indicates “strongly agree.”

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
1. To complete the questions above, using this search engine was very frustrating.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
2. To complete the questions above, using this search engine took too much time.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
3. To complete the questions above, using this search engine required too much effort.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
4. To complete the questions above, using this search engine was too complex.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
5. To complete the questions above, using this search engine was easy.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
6. It is easy for me to move to the target page.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
7. It is convenient for me to look for detailed information.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7

Please indicate the degree to which you would agree with the following statements on the format that you just used by choosing a number from 1 to 7, where 1 indicates “strongly disagree” and 7 indicates “strongly agree.”

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
1. There are definitely times that I would like to use this search engine.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
2. I would like this search engine available for my use all the time.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
3. I found this search engine useful.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
4. I found this search engine confusing to use.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
5. I liked this search engine. .	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7

For the below questions, you will need to go to more than one URL to find information to answer the question. Record all the URLs that you reviewed to answer the question.

Time started: __________

Question 6: Describe the similarities and differences in the political views in education of Hilary Clinton, Barack Obama, and John McCain.

Answer

URL(s)

Question 7: Imagine you are writing a paper on music. Describe how the genre of "hip hop" in 2008 has influenced our culture

Answer

URL(s)

Time completed: __________

Please indicate the degree to which you would agree with the following statements on the given information format that you used to complete the above tasks by choosing a number from 1 to 7, where 1 indicates “strongly disagree” and 7 indicates “strongly agree.”

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
1. To complete the questions above, using this search engine was very frustrating.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
2. To complete the questions above, using this search engine took too much time.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
3. To complete the questions above, using this search engine requires too much effort.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
4. To complete the questions above, using this search engine was too complex.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
5. To complete the questions above, using this search engine was easy.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
6. It is easy for me to move to the target page.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
7. It is convenient for me to look for detailed information.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7

Please indicate the degree to which you would agree with the following statements on the format that you just used by choosing a number from 1 to 7, where 1 indicates “strongly disagree” and 7 indicates “strongly agree.

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
1. There are definitely times that I would like to use this search engine.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
2. I would like this search engine available for my use all the time.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
3. I found this search engine useful.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
4. I found this search engine confusing to use.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
5. I liked this search engine. .	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7

APPENDIX F – Normality Tests and Plots

Descriptives

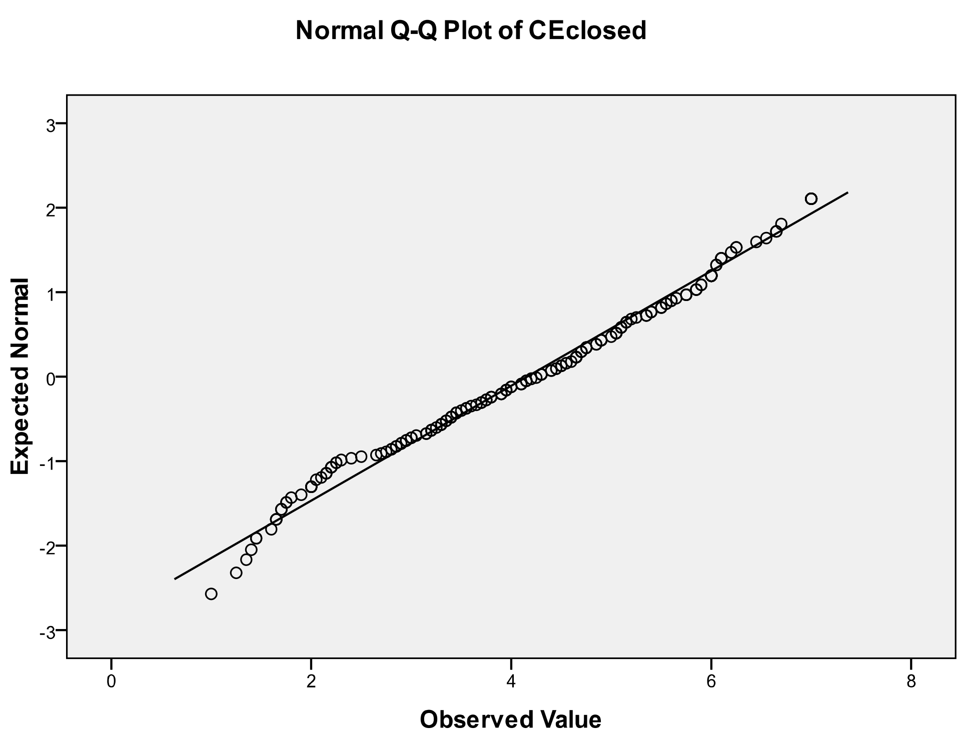
			Statistic	Std. Error
CEclosed	Mean		4.1584	.10476
	95% Confidence Interval for Mean	Lower Bound	3.9518	
		Upper Bound	4.3650	
	5% Trimmed Mean		4.1592	
	Median		4.3000	
	Variance		2.162	
	Std. Deviation		1.47039	
	Minimum		1.00	
	Maximum		7.00	
	Range		6.00	
	Interquartile Range		2.03	
	Skewness		-.116	.173
	Kurtosis		-.796	.345
CEopen	Mean		3.6289	.11774
	95% Confidence Interval for Mean	Lower Bound	3.3967	
		Upper Bound	3.8611	
	5% Trimmed Mean		3.6538	
	Median		3.6000	
	Variance		2.731	
	Std. Deviation		1.65256	
	Minimum		.00	
	Maximum		7.00	
	Range		7.00	
	Interquartile Range		2.25	
	Skewness		-.164	.173
	Kurtosis		-.242	.345

Tests of Normality

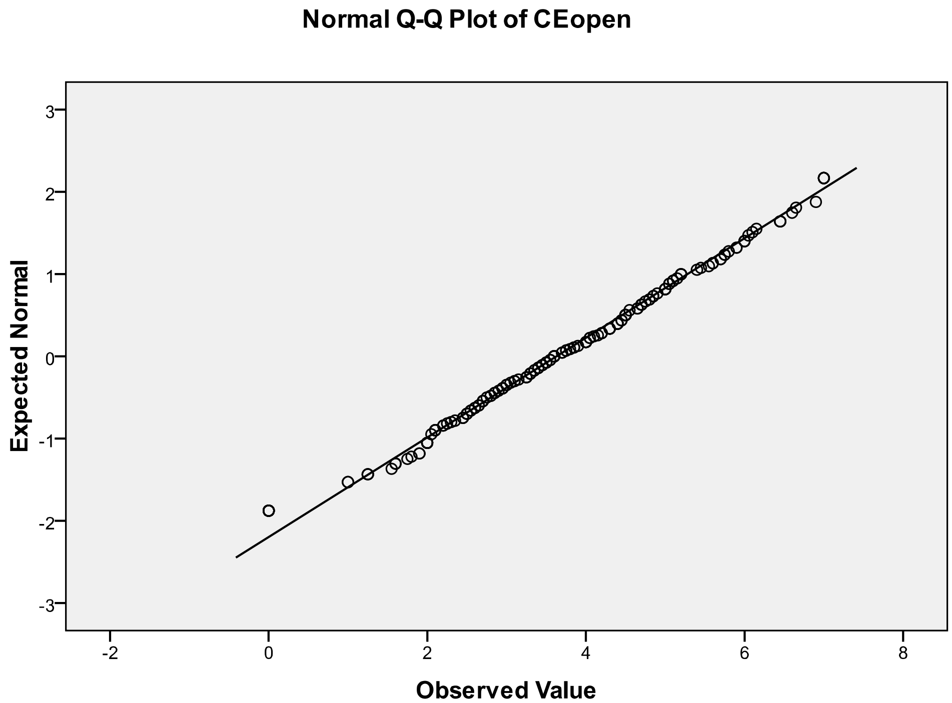
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
CEclosed	.060	197	.079	.979	197	.005
CEopen	.042	197	.200*	.983	197	.016

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.



Normality Plot for CEclosed (Closed-ended Cognitive Effort)



Normality Plot for CEopen (Open-ended Cognitive Effort)

APPENDIX G – Statistical analyses for Concatenated Formats on Cognitive effort for Closed-ended Tasks

Tests of Between-Subjects Effects of Concatenated Formats on Cognitive Effort for Closed-ended Tasks

Dependent Variable:CEclosed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
cognitivestyle	13.477	1	13.477	6.896	.009
gpcat	38.496	1	38.496	19.699	.000
Sex	.223	1	.223	.114	.736

a. R Squared = .115 (Adjusted R Squared = .096)

Estimates of Mean Cognitive Effort for Closed-ended Tasks

Dependent Variable:CEclosed

gpcat	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	3.714[a]	.141	3.436	3.992
2	4.605[a]	.142	4.325	4.886

a. Covariates appearing in the model are evaluated at the following values: cognitivestyle = .74.

APPENDIX H – Statistical analyses for Concatenated Formats on Cognitive effort for Open-ended Tasks

Tests of Between-Subjects Effects of Concatenated Formats on Cognitive Effort for Open-ended Tasks

Dependent Variable:CEopen

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
cognitivestyle	14.316	1	14.316	5.321	.022
gpcat	4.561	1	4.561	1.695	.194
Sex	.376	1	.376	.140	.709

a. R Squared = .035 (Adjusted R Squared = .015)

Estimates of Mean Cognitive Effort for Open-ended Tasks

Dependent Variable:CEopen

gpcat	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	3.480[a]	.166	3.153	3.806
2	3.786[a]	.167	3.457	4.115

a. Covariates appearing in the model are evaluated at the following values: cognitivestyle = .74.

APPENDIX I - Matrix for the Mediator Variables for Closed-ended Tasks

```
Dependent, Independent, and Proposed Mediator Variables:
DV =   SATclose
IV =   CEclosed
MEDS = SCOREclo
       TIMEclos

Statistical Controls:
CONTROL= Age
         English
         experien
         UseInter
         writingq
         motivati

Sample size
        197

IV to Mediators (a paths)
             Coeff        se         t         p
SCOREclo    -.1927     .0523   -3.6856     .0003
TIMEclos   87.1848  158.4772     .5501     .5829

Direct Effects of Mediators on DV (b paths)
             Coeff        se         t         p
SCOREclo    -.0043     .0637    -.0669     .9468
TIMEclos     .0000     .0000   -1.9223     .0561

Total Effect of IV on DV (c path)
             Coeff        se         t         p
CEclosed    -.8406     .0460  -18.2839     .0000

Direct Effect of IV on DV (c-prime path)
             Coeff        se         t         p
CEclosed    -.8379     .0474  -17.6747     .0000

Partial Effect of Control Variables on DV
             Coeff        se         t         p
Age         -.0157     .0103   -1.5264     .1286
English      .2894     .1654    1.7494     .0819
experien    -.3897     .1422   -2.7408     .0067
UseInter    -.0954     .0670   -1.4247     .1559
writingq     .0767     .0974     .7876     .4319
motivati     .0609     .0742     .8201     .4132

Model Summary for DV Model
      R-sq  Adj R-sq         F       df1       df2         p
     .6682     .6522   41.8443    9.0000  187.0000     .0000

*****************************************************************
```

```
BOOTSTRAP RESULTS FOR INDIRECT EFFECTS

Indirect Effects of IV on DV through Proposed Mediators
(ab paths)
              Data      boot      Bias        SE
TOTAL       -.0027    -.0023     .0004     .0136
SCOREclo     .0008    -.0003    -.0011     .0114
TIMEclos    -.0035    -.0020     .0015     .0073
C1           .0043     .0017    -.0026     .0135

Bias Corrected and Accelerated Confidence Intervals
             Lower     Upper
TOTAL       -.0323     .0237
SCOREclo    -.0222     .0269
TIMEclos    -.0337     .0040
C1          -.0204     .0341

*****************************************************************

Level of Confidence for Confidence Intervals:
  95

Number of Bootstrap Resamples:
  1000

*****************************************************************

  INDIRECT EFFECT contrast DEFINITIONS: Ind_Eff1 MINUS Ind_Eff2


  contrast  IndEff_1  IndEff_2
  C1        SCOREclo  TIMEclos

------ END MATRIX -----
```

APPENDIX J - Matrix for Mediator Variables for Open-ended Tasks

```
Dependent, Independent, and Proposed Mediator Variables:
DV =   SATopen
IV =   CEopen
MEDS = SCOREope
       TIMEopen

Statistical Controls:
CONTROL= Age
         English
         experien
         UseInter
         writingq
         motivati

Sample size
        197

IV to Mediators (a paths)
             Coeff        se         t         p
SCOREope    -.0017     .0047    -.3689     .7126
TIMEopen  161.0958   52.9013    3.0452     .0027

Direct Effects of Mediators on DV (b paths)
             Coeff        se         t         p
SCOREope    4.5854    1.0333    4.4377     .0000
TIMEopen     .0005     .0001    4.9595     .0000

Total Effect of IV on DV (c path)
           Coeff        se         t         p
CEopen    -.3069     .0737   -4.1674     .0000

Direct Effect of IV on DV (c-prime path)
           Coeff        se         t         p
CEopen    -.3723     .0683   -5.4482     .0000

Partial Effect of Control Variables on DV
             Coeff        se         t         p
Age         -.0242     .0168   -1.4350     .1530
English      .6538     .2746    2.3807     .0183
experien    -.7247     .2338   -3.0997     .0022
UseInter    -.1170     .1105   -1.0593     .2908
writingq    -.0449     .1610    -.2791     .7805
motivati     .0553     .1228     .4505     .6529

Model Summary for DV Model
      R-sq  Adj R-sq         F       df1       df2         p
     .2950     .2611    8.6935    9.0000  187.0000     .0000

*****************************************************************

           BOOTSTRAP RESULTS FOR INDIRECT EFFECTS
```

```
      Indirect Effects of IV on DV through Proposed Mediators (ab
paths)
                    Data       boot       Bias         SE
      TOTAL        .0653      .0669      .0016      .0488
      SCOREope    -.0079     -.0079      .0000      .0246
      TIMEopen     .0733      .0748      .0015      .0442
      C1          -.0812     -.0827     -.0015      .0523

      Bias Corrected and Accelerated Confidence Intervals
                   Lower      Upper
      TOTAL        -.0169      .1812
      SCOREope    -.0558      .0388
      TIMEopen     .0129      .2214
      C1          -.2291     -.0054

      *****************************************************************

      Level of Confidence for Confidence Intervals:
        95

      Number of Bootstrap Resamples:
        1000

      *****************************************************************

        INDIRECT EFFECT contrast DEFINITIONS: Ind_Eff1 MINUS Ind_Eff2


        contrast  IndEff_1  IndEff_2
        C1        SCOREope  TIMEopen

      ------ END MATRIX -----
```