

Bogeum Choi. The relationship between task complexity, working memory and information search. A Master's Paper for the M.S. in I.S degree. April, 2018. 59 pages.  
Advisor: Robert Capra

In recent years, there has been growing interest in better understanding human actors in human information seeking behavior studies. Although a number of studies have been conducted to explore users' individual differences in search behavior, there have been few studies taking both a theoretical and empirical approach to the relationship between users' cognitive ability, task complexity and search interactions. The study presented in this paper evaluated the effect of task complexity and working memory in human information searching behavior. Twenty-four participants from a non-college-bound adolescents sample (ages 18-50+) performed two search tasks of varying levels of complexity and were administered measures of working memory. ANOVA tests revealed three important trends: (1) task complexity had a significant main effect on users' perceptions about the task (i.e., temporal demand and level of satisfaction with time spent on the task), (2) working memory capacity had a significant main effect on users' search behavior (i.e., queries, clicks, time until 1st click and time between search activities), and (3) a significant interaction effect was found for several search interaction measures (i.e., queries, clicks and time between search activities) and perceived level of temporal demand. Specifically, participants with high working memory capacity carried out more search activities at a faster pace and experienced less temporal workload. Taken together, these results suggest that task complexity and working memory capacity can have effects on users' search behavior as well as their perceptions about the search experience.

#### Headings:

Information retrieval

User studies

Cognitive ability

Working memory

Task complexity

Locus of control

Search behavior

THE RELATIONSHIP BETWEEN TASK COMPLEXITY, WORKING MEMORY  
AND INFORMATION SEARCH

by  
Bogeum Choi

A Master's paper submitted to the faculty  
of the School of Information and Library Science  
of the University of North Carolina at Chapel Hill  
in partial fulfillment of the requirements  
for the degree of Master of Science in  
Information Science.

Chapel Hill, North Carolina

April 2018

Approved by

---

Robert Capra

## Table of Contents

1. Introduction.....	2
2. Background.....	4
2.1 Cognitive ability.....	4
2.1.1 Working memory .....	6
2.2 Task complexity.....	9
2.3 Workload.....	10
2.4 Locus of control .....	12
3. Research questions.....	16
4. Methods.....	17
4.1 Participants.....	18
4.2 Procedures.....	18
4.3 Working memory task.....	20
4.4 Search tasks.....	22
4.5 Post-task survey .....	24
4.6 Search behavior measures.....	25
4.7 Search outcome measures.....	26
5. Results.....	30
5.1 Working memory .....	30
5.2 Effects on users' perception.....	31
5.3 Effects on users' search behavior .....	34
6. Discussion.....	38
7. Conclusion .....	42
8. Acknowledgements.....	44
9. Reference .....	45
Appendix 1. Consent form.....	50
Appendix 2. General instructions page.....	53
Appendix 3. Search tasks .....	54
Appendix 4. Post-task survey .....	55

## **1. Introduction**

Information seeking behavior can be considered as an assemblage of cognitive activities. That being said, it is key to understand the extent to which variations in cognitive abilities can impact one's search behaviors, perceptions, and outcomes. In the field of behavior analysis, there have been a substantial number of studies focusing on consistent differences in the level of performance among individuals and cognitive abilities have received substantial attention in this regard (Kuncel, Hazlette & Ones, 2004; Rohde & Thompson, 2007; Schmitt, 2014; Ruffing, Wach, Spinath, Brunken & Karbach, 2015). However, not so much attention has been drawn to individual variation in cognitive abilities in information behavior studies. In an educational setting, individual differences in performance are among the most salient aspects of behavior. Some people learn and understand the complex material with relative ease, whereas others labor to succeed. Learning and searching share core areas of commonality in that they both require interaction with texts and becoming informed as a result of the actions, which means search behavior may also be under the influence of such individual differences. Among those individual differences factoring into the dynamics of human information behavior, this study focuses mainly on cognitive abilities, working memory in particular. In this study, I also manipulated complexity of task as an attempt to see how individuals with varying degree of cognitive ability accommodate search behavior to cope with challenges given by a task.

The study can contribute to the body of knowledge of human information behavior with emphasis on human factors in the dynamics between users, task, and system. Findings from this study can also help us in making interactive information systems more attuned to individual differences and the variance of tasks. Furthermore, it may lead us to develop supplementary tools that allow users to utilize the system in a way that works best for them, supporting both users and system.

## **2. Background**

This section describes four areas of prior work related to this study: cognitive abilities, task complexity, workload and locus of control.

### **2.1 Cognitive ability**

Cognitive science sees the individual as a processor of information, in much the same way that a computer takes in information and follows a process to produce an output. This basic assumption of cognitive science led to the emergence of the information processing model as a framework for studying human mind. To envision how this information processing approach can provide a good fit to the mental model of information seeking behavior, I include Figure 1 below. Figure 1 is developed from the original information processing model proposed by Atkinson and Shiffrin (1968) which is also known as a multi-store model of memory.

Marchionini (1995) produced a model of the information-seeking process for computer-based searches. Among the eight phases that consist the model are recognizing information problem, formulating a query by which to fit the information need, examining results and extracting information from the results. Kulhthau (1991) also argued cognitive actions such as locating relevant information about the general topic, reading to become informed and relating new information to what is already know is required in the “exploration” stage of her information search process model. These approaches suggest that it is key to examine the cognitive process (i.e., perception,

attention, encoding, rehearsal and retrieval) embedded in each of the main activities in a search process as a low-level action.

This approach can be a good way to see view search process in that it allows us to clearly distinguish the cognitive processes embedded in a search process, which include perception, attention, encoding, rehearsal, and retrieval.

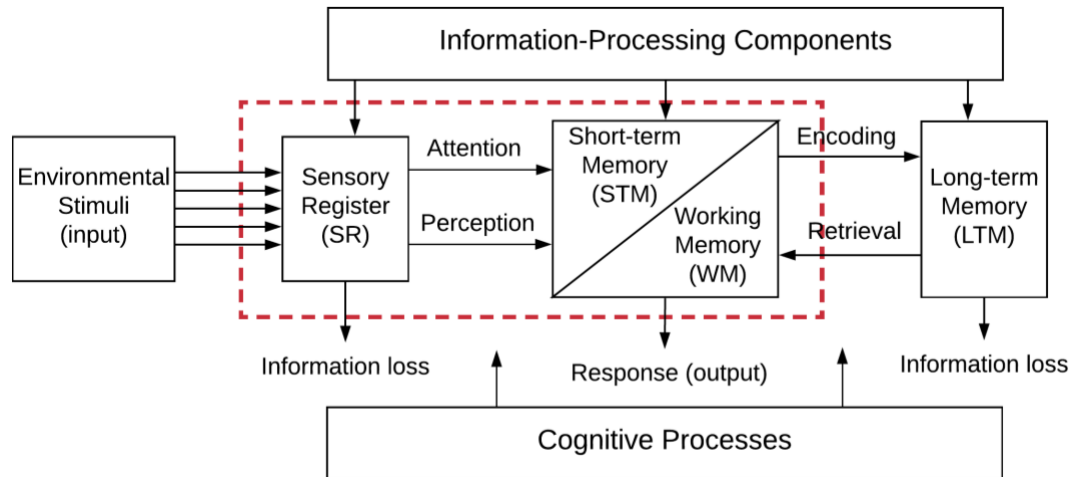


Figure 1. Information processing model (Atkinson & Shiffrin, 1968)

A psychological approach to information studies has been taking up its position as a dimension of human information behavior studies, focusing on searching behavior (Wilson, 2000). In early computing and interactive information retrieval (IIR), cognitive ability was introduced as a factor that shapes the interaction between users and system. Early studies investigated the influence of cognitive abilities in text editing systems, hierarchical file systems and databases, and general computer-based software programs. Cognitive abilities were found to be causal factors of individual differences in task performance, accuracy, and navigation. For example, those with higher visual memory skill completed text editing task faster and with fewer errors (Egan, Bowers & Gomez,

1982; Gomez, Egan, Wheeler, Sharma & Gruchacz, 1983). Gomez, Egan and Bowers (1986) also found that those who showed higher associative memory performance made fewer first-try errors and completed an editing task faster. A few other cognitive abilities such as perceptual speed (Salthouse & Babcock, 1991; Fisk & Warr, 1996; Brennan, Kelly & Arguello, 2014; Turpin, Kelly & Arguello, 2016) or working memory (Gwizdka, 2009; MacFarlane, Albrair, Marshall & Buchanan, 2012; Gwizdka, 2013; Gwizdka, 2017) have been found to serve as factors associating with searching behavior.

### **2.1.1 Working memory**

One of the core abilities that holds great influence throughout the whole cognitive process in the information processing model is working memory. Working memory refers to a cognitive ability responsible for providing access to information required for ongoing cognitive processes. Working memory is critical for making sense of anything that unfolds over time and that requires holding in mind what happened earlier and relating that to what comes later. For example, doing any math in your head requires working memory, as does incorporating new information into your thinking or action plans (updating), considering alternatives, and mentally relating information to see relations between items or ideas (Diamond, 2013).

Miller (1956) made an early argument that there is a limitation that is imposed on our ability to process information. The length of chunked items that an individual can maintain in memory-storage is called memory span and this limited capacity has a profound effect on human information processing and interaction with the external world, including the interaction with information. In this paper, working memory capacity



(WMC) is used to refer to an individual differences construct reflecting the limited capacity of a person's working memory.

Working memory has been studied as a factor that can impact search behavior in IR literature. Gwizdka (2009) explored the relationships between selected tasks, cognitive abilities (i.e., working memory, verbal closure) and search result interfaces. The study found that searchers with higher cognitive abilities were faster in the interface where they were provided an overview interface (vs. list interface with no overview of results shown) during less demanding tasks. Searchers with higher cognitive abilities also exerted more search effort in the interface where tags were presented in a form of a list without an overview while performing more demanding tasks. These findings suggest the importance of considering cognitive ability such as working memory in the design of search results presentation.

As a follow-up to the previous study, Gwizdka (2013) explored the effects of task complexity and working memory span on participants searching a collection of social bookmarks related to travel, sightseeing, and shopping, under two different interface conditions. The study reported the interaction between memory span and task complexity for task duration and query behaviors. On simple tasks, participants with higher memory span spent less time, issued fewer queries, and opened fewer documents than participants with lower memory span. On more complex tasks, however, participants with higher memory span spent more time, issued more queries, and opened more documents. They also performed cognitive actions (i.e., issuing and reformulating queries, opening a document) faster than the low working memory group.

Gwidka (2017) also investigated the effects of working memory on search effort. The findings show that searchers with higher working memory perform more actions than searchers with lower working memory, confirming results from previous work. Also, it has shown that participants with higher working memory spent significantly longer time on completing a task, particularly on reading search results than participants with lower working memory. The study also noted a trend that higher working memory participants showed their efforts across different phases of search in a more consistent manner than lower working memory participants.

In another empirical study MacFarlane and his colleagues (MacFarlane, Albrair, Marshall & Buchaman, 2012) studied the impact of working memory on information searching by university students using TREC collections and topics. They recruited eight dyslexic students and eight non-dyslexic students to compare two different user groups. There is a compelling evidence that cognitive deficit that causes dyslexia lies in the impaired phonological working memory. The authors assumed that reading ability of users would show different search behaviors given that reading is one of the key activities required for searching. They found that the number of documents marked irrelevant was significantly correlated with a measure of working memory. Participants with higher working memory viewed more documents in total during the task completion and they marked more document as non-relevant than participants with lower working memory (i.e., the percentage of documents judged non-relevant as well as the absolute number of the documents), which demonstrates the impact of impaired working memory on search behavior. Findings of these studies suggest working memory can play an important role in search.

## 2.2 Task complexity

Previous works have found that task complexity affects search behavior (Jansen, Booth, & Smith, 2009; Wu, Kelly, Edwards & Arguello, 2012) and post-task assessments of task difficulty (Arguello, 2014; Wu, Kelly, Edwards & Arguello, 2012). Various approaches to viewing task complexity were taken by different researchers. Early work by Wood (1986) regarded complexity as depending on the number of desired outcomes, the number of actions required to produce the outcomes, and the quality of the information cues processed during the task. Additionally, Campbell (1986) mentioned that task complexity can be measured by the number of required outcomes, the number of alternative paths to the outcomes, the level of uncertainty regarding the paths, and the degree of interdependence between subparts consisting a task. He also found that task complexity has a significant impact on participants' perception of their task performance as well as objective measures of task performance.

Gwizdka (2013) examined user behavior on information search tasks at two levels of complexity (i.e., simple and complex) and reported the significant main effect of task complexity on time and search effort as well as the length of eye fixation. Walhout, Oomen, Jarodzka & Brand-Gruwel (2017) explored the effect of search task complexity on search query formulation, evaluation of search results, and task performance. The study showed that an increase in task complexity results in more search queries, more time spent on formulating queries, and higher consideration for results from SERPs (i.e., search engine result pages). Brennan, Kelly, and Arguello (2014) conducted an empirical study to evaluate the relationship between cognitive ability, task complexity and information search behavior and found a significant relationship between task complexity

and workload, and task complexity and search behaviors. While the interactions were not significant, the differences were found to be more pronounced for more complex tasks.

Prior studies showed an evidence of the effect of task complexity and it's reasonable to assume that complex task will invite a user to more challenges in information search. In this study, I used task complexity as another independent variable, hoping that it will give me more room to explore individual differences in users' search behavior and perception by causing a different quantity of demands on users.

### **2.3 Workload**

The concept of workload and mental workload was emerging as an important aspect to be considered not only in the evaluation of professional task performance but also in the domain of scientific research (Silva, 2014). A theoretical approach to mental workload has been associated with the need for a proper understanding of the interaction between human-machine systems, the advantages and limitations underlying this interaction and also the outcomes as a result of the interaction. Mental workload began to be associated with studies on safety and effectiveness of operator performance in many organizational contexts where the concept was primarily related to the limit and amount of mental effort, and tasks allowing to maintain an adequate level of performance.

De Waard (1996) pointed out that mental workload not only reflects external demands placed by a task but also inherently reports a specific person's situation. Kahneman (2002) mentioned that workload can be defined by the amount of resource required by a set of concurrent tasks, as well as by the use of resources needed to perform them. In an attempt to minimize the confusion of the usage of the term, I have done my

best to clearly describe the construct I was measuring by distinguishing two main aspects of mental workload. The amount of workload an individual experiences during the task is a result of that individual's capability (i.e., defined by cognitive abilities), task demand (and features) and the person's situation (Jex, 1988).

The measurement of workload has been used to understand search tasks in different ways by IIR researchers. Haapalainen, et al. (2010) measured mental demand to verify the levels of workload imposed to participants in their study. They combined workload scores with task completion time to conduct a manipulation check of task difficulty. Workload has also been incorporated to see other effects it has on users' experiences. For example, it was found that text vs. visual based interface had a significant effect on workload; reported mental load for the visual query interface was significantly lower than that for the text-based approach. (Speier et al., 2003). Kelton and Pennington (2012) found that information presentation format (i.e., hyperlink vs. paper) can affect users' decision-making performance in terms of the amount of mental effort put on the task. In another study, Gao (2011) investigated users' motivation, performance, and workload when they use tagging interface (as opposed to categorization interface) to organize personal information system. He found that participants who used a tagging interface reported a significantly higher level of mental demand and frustration when performing organizational tasks.

In this study, I used the NASA-Task Load Index (TLX) to measure workload, which measures the mental, physical, and temporal demands imposed on individuals by a work task along with individual's evaluations of their performance, effort and experienced frustration (Hart & Staveland, 1988).

## 2.4 Locus of control

Locus of control, in social psychology, refers to the extent to which individuals believe that they can control events and outcomes in their own lives (Rotter, 1966), as opposed to external forces beyond their control. The idea was first introduced by Julian B. Rotter (1954) who conceptualized a person's loci (plural of locus; Latin for place or location) being on a continuum from internal to external. Individuals with strong internal locus of control believe that the responsibility for the outcome of life activities lies with themselves. Internals believe that success or failure is due to their own efforts. On the other hand, externals believe that the reinforcers (i.e., stimuli given as a consequence of one's action, in other words, an outcome of one's behavior) in life are controlled by luck, chance, or powerful others. Therefore, they see little impact of their own efforts on the amount of reinforcement they receive.

It is documented that perceived causal attribution influences changes in goal expectancy thus is presumed to guide motivated behavior (1985, Weiner). Within social psychology, the work of Heider (1958), deCharms (1968; 1976), Jones and Nisbett (1972) and others have focused upon a causal explanation of events as major determinants of the behavior exhibited during those events. To give a real-world example, if students either attribute their academic success or failure to having a bad day, or unfair grading procedures on their teacher's part, they can be said to have a more external locus of control. Since they attribute both their successes and failures to luck or chance, they tend to lack persistence and not have very high levels of expectation. Lei (2009) found that learners can be most persistent at academic tasks if they attribute their

academic success to internal factors that they have control over, such as effort. It leads to the conclusion that for students to be able to be persistent in academic activities, they need to believe that they are competent and that by working hard they can be successful.

Information seeking behavior can be defined as the purposeful seeking for information as a consequence of a need to satisfy some goal (Wilson, 2000). In the course of seeking, the individual interacts with either manual or computer-based systems. Here, we can identify three main entities involved in the process; human, system and information need that primarily relates to the context where the interaction between human user and system takes place. In interactive information retrieval setting information need is usually imposed on a user by a task at hand, and the notion of task completion can go parallel to goal attainment in an achievement-related setting. I thought the construct may provide a good approach to better understand users' perception for their own search performance, and furthermore, I expected an explanatory power of this concept in explaining patterned variations of users' search behavior as well as achieved search outcome.

In order to measure the locus of control in the context of human information seeking behavior, I adapted the multi-dimensional multi-attributional causality scale (MMCS) that was introduced by Lefcourt, von Baeyer, Ware, and Cox (1979). The original MMCS is composed of two parts, each of which deal with causal beliefs about their respective areas; achievement and affiliation. It comprises 48 items, 24 dealing with achievement, 24 with affiliation. Affiliation locus of control assumes interpersonal situations and since I'm dealing with an achievement-related setting, I excluded the affiliation part. The scale I modified is the fifth revision with improved reliability figures.

Half of each set concern success and half failures. Within each 12-item set, there are four attributions composed of three items each. These consists of ability and effort as the internal attributions (also known as dispositional attributions), and luck and context as the external attributions (also known as situational attributions).

Lefcourt et al. (1979) found the percentage of time that subjects showed to complete a given task (i.e., anagram procedure) was significantly related to achievement locus of control scores as well as the interaction effects with difficulty stage of the task (simple - intermediate - difficult - unsolvable). In brief, the result suggested that the more internal were the subject's expectancies for achievement, the more likely the subjects were to exhibit signs of discomfort during the task, particularly during the more difficult sections of that task when failure became increasingly probable. This suggests that individuals' belief in their control of reinforcement can specifically affect their reaction to hardships associated with a task.

Earlier work by Dweck and her colleagues (Dweck, Davidson, Nelson & Enna, 1979; Dweck & Reppucci, 1973) have also demonstrated that difficulties in learning that result from the tendency to give up in the face of failure can be accounted for by attributions of ability as opposed to effort. They found that a child having difficulty with arithmetic is more apt to stop trying and give up if he or she construes failures as evidence of his or her lack of ability. On the other hand, if attributions are made to an effort, persistence despite failure turned out to be more likely. Again, given that the scale has been successfully used to predict things including academic success and one's reaction to hardships, I assumed that the locus of control manifested by a user may reveal



interesting dynamics of users' perception and behavior while they are completing a search task with varying complexity.

### **3. Research questions**

Based on the review of the literature, I posit several research questions about the effects of working memory and task complexity on search experience. My manipulations and variables are described in the next section.

RQ 1: How does task complexity affect users' a) perception, b) search behavior, and c) search outcome?

RQ 2: How does variance of the working memory impact users' a) perception, b) search behavior, and c) search outcome?

RQ 3: Are there interaction effects of task complexity and working memory on a) perception, b) search behavior, and c) search outcome?

#### 4. Methods

A controlled experiment study was conducted to investigate my three main research questions (R1-R3). The study involves two dependent variables; working memory capacity and task complexity, each of which has two levels and looks at three categories of dependent variables; search behavior, search outcome and users' perception. The overall study design with variables of my interest is presented in Figure 2.

Working memory capacity was a between-subject variable that consisted of two groups: high and low. Participants were divided up into two groups (i.e., high and low) according to their memory span for operation task split at the median. Task complexity was manipulated as a within-subject variable and consisted of two levels: simple and complex. All participants were exposed to two search tasks with each task associated with a different level of complexity.

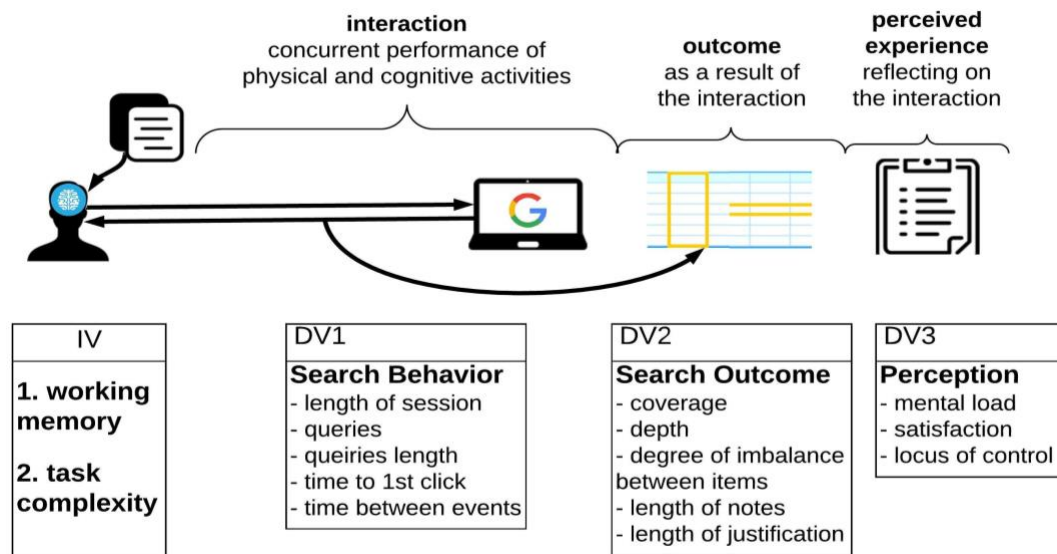


Figure 2. Variables of interest in the study setting

## **4.1 Participants**

To increase the chances of obtaining a sample with a wider range of cognitive abilities, I recruited twenty-four participants from university employees, rather than relying on to university students. Additionally, a number of laboratory user studies in IIR have been conducted with university students, and I wanted to contribute to the research of which findings are more generalizable to the general public. A total 24 participants consisting of 18 females and 7 females (aged 18–late 50+ years) completed this study. The study was conducted in Interactive Information System Lab (IISL) at the University of North Carolina at Chapel Hill.

## **4.2 Procedures**

The study protocol was as follows. Upon arrival to the session, the researcher briefed participants on the goals of the study and gave an overview of the study session (i.e., they were informed that there are two search tasks, a post-task survey for each task and memory task to perform). Participants then were given an informed consent form (see Appendix 1) that acknowledged: that participation is voluntary, that they may cease at any time with no penalty and that no personal identifiers will be collected during the session. Participants were randomly assigned to treatments and given a unique participant ID (i.e., a number between 1 and 24). After they agreed to participate the study, participants were directed to a custom-built search engine and logged into the system by entering their participant ID. The system used in this study was provided by Jaime Arguello. The system was a modified version of the system used in one of his studies

(Turpin, Kelly & Arguello, 2016). The system used the Bing Web Search API to retrieve results from the open web and allowed participants to issue queries, click and view results, and navigate away from the landing page. The system also logs user interactions on the search engine result page including scroll and mouse movements as well as timestamp mapped to each of those behavior using JavaScript and AJAX.

After they fill out a brief demographic questionnaire, the researcher pulled up a Microsoft Word document on the right side of the desktop screen (see Appendix 2). General instructions for task completion were shown on the first page of the document which was followed by a task description and a search outcome table for the two tasks of varying complexity. The instructions for the memory task was displayed on the last page. Participants were encouraged to read along the written instructions while the moderator was explaining it to them. An example of a comparative task (i.e., a type of task used in the study) was included on the first page with an empty table to ensure that participants clearly understand the characteristics of the task. The detailed information about a comparative task is provided in Section 4.4. To help participants remember the task and to make it easier for participants to fill out a given table while searching for information, the Word document remained visible at the right side of the desktop screen during the whole search session.

Participants completed two search tasks of different levels of complexity. The order of search tasks was rotated across participants. For each task, participants were asked to create a response by either typing notes or copying and pasting information they found relevant (i.e., URL of a webpage or text from useful web pages) into a given table in a document. The task instructions included the following quotations: “Your goal is to

find as much information as you think is needed to address each cell in the table. ... Time that takes for you to complete is task is not important; but for the sake of your time, you will be given twenty minutes per task. You don't have to use up the whole time. You can stop searching for information when you feel like you have the adequate amount of information to make a decision.”

After performing two search tasks, they completed a post-task which contained items about workload, perceived satisfaction with search performance, and locus of control questionnaire (see Appendix 4). Participants then conducted a working memory span task (i.e., OSPAN) which lasted 10 to 15 minutes. At the end of the study, the participant received the compensation of \$15 and signed on a paper copy of a receipt. The study took approximately 60 minutes in total.

### **4.3 Working memory task**

In their extensive methodological review on working memory (WM) span tasks, Conway et al. (2005) argued that working memory system would be unlikely to evolve for the sole purpose of allowing a human to store or rehearse information (e.g., phone number) while it was doing nothing else. Rather, a more adaptive system would allow the organism to keep task-relevant information active and accessible in memory during the execution of other cognitive tasks. It explains the purpose of complex working memory span tasks (i.e., reading span task, operation span task, and counting task) where a participant is required not only to store and rehearse an information but also to process additional information that occurs simultaneously (Case, Kurland & Goldberg, 1982; Turner & Engle, 1989).

In the light of Conway's argument, in this study, Turner and Engle (1989)'s operation-word-span task (OSPAN) was used to measure working memory span. In OSPAN task, participants are asked to read out loud and verify a simple math problem such as "Is  $(4/2) - 1 = 1$ ?" and then read a word after the operation such as "snow". After a series of problems and words have been presented, the participants recall the words that followed each operation. Three trials of each set size (i.e., list length is from 2 to 6) are presented, with the order of set size varying randomly, so that participants cannot predict the number of items they will see in the next sequence. A participant's operation span is calculated by adding up the length of all word lists perfectly recalled in order. Additionally, in order to ensure that participants are not trading off between solving the operations and remembering the words, an 85% accuracy criterion on the math operations is required.

The task has demonstrated good reliability and validity (Conway et al., 2002; Engle, Tuholski, Laughlin & Conway, 1999; Klein & Fiss, 1999). For example, with a sample size of 236, Kane et al. (2004) observed coefficient alphas of .80 for operation span, which indicates that subjects who responded with the correct answer for one set of span stimuli in the task (e.g., equation word pairs) tended to respond with the correct on the others (and vice versa). Evidence also suggested that the task is reliable in the sense that the rank order of span scores is stable across time. Test-retest correlations of .70-.80 have been observed for operation span task over weeks, and over 3 months (Klein & Fiss, 1999). The task was operated using the CogLab 2.0 CD from a cognitive lab textbook (Francis, 2003).

#### 4.4 Search tasks

In this study, I used comparative search tasks. A comparative search task is a type of task in which one has to compare a set of items along different dimensions. Tasks were designed at two levels of complexity and the complexity level was manipulated by the numbers of items and dimensions that had to be considered during a search. Simple tasks required a comparison between two items along two dimensions; whereas, complex task involved four items with four dimensions. This increased the level of complexity by creating multiple interdependent paths to the task outcome. A task scenario was constructed to present realistic situations to provide participants with the search context and the basis for relevance judgments. In all tasks, participants were motivated to look for information for their friend. At the end of the task, they were asked to make a decision to choose one item among given items and generate the answer in the format of giving a suggestion to their friend. An example of task description is displayed in Table 2. Tasks were presented to participants in a counter-balanced order in terms of both topic and levels of complexity.

Table 2. Assigned search task example (items X dimensions)

Topic	Level of complexity	Scenario	Description
1	Simple (2X2)	A friend of yours has recently decided to get a dog for companionship. Your friend works during the day and lives in an apartment complex.	How do a) Pug and b) Bichon Frise dog breeds differ as a choice for your friend in terms of a) its ability to be left alone and b) need for outdoor activity?
	Complex (4X4)	Now you are trying to help your friend make a decision.	How do a) Pug, b) Bichon Frise, c) Beagles and d) Golden Retriever differ as a choice for your friend in terms of a) its ability to be left alone, b) need for outdoor activity, c) how well they do in small space and d) temperament?



There are two main purposes for using a comparative task in this study. First of all, cognitive activities that comparative task involves explicitly require the role of working memory. For instance, in order to perform a task, a participant has to first recognize different items, search for relevant information that corresponds to a specific condition (e.g., cost of nicotine replacement therapy), then compare a piece of information to another. This requires processing new information, while retaining prior information. Besides its explicit reflection of the role of working memory, I was also able to manipulate the level of complexity in a fairly straightforward way by simply adjusting the number of components (i.e., item and dimension) that consist a task. An individual's working memory capacity is often represented in the number of discrete units of information over which one can successively distribute one's attention and maintain in an active state. Therefore, it seems reasonable that increasing the number of elements to consider could increase the memory load.

Secondly, search outcome of a comparative task can be generated in the form of a table with columns and rows, which can give me a visual representation of distribution of information (i.e., degree of balance or imbalance across the items). As such, this two-dimensional matrix structure enables an analysis of measures such as coverage, depth, and degree of imbalance of search outcome. These measures will be discussed in more detail in section, 4.7 Search outcome measures.

#### 4.5 Post-task survey

Perceptions are the basic means by which people experience the world and build a worldview to explain those experiences. As opposed to sensations serving as direct sensory stimuli, perceptions are the ways we interpret those sensations to make sense of what we are sensing. This paper will also examine users' perceptions of their own search experience as a main category of dependent variable.

I used a questionnaire (see Appendix 4) with 17 items in order to measure (1) workload (2) satisfaction and (3) locus of control for each task. For the work load questions, I selected four items from six items of the NASA Task Load Index (NASA-TLX). I took out an item that measures physical demand, judging that physical activities involved in performing web search tasks will not say much about the variables of interest of this study. I also didn't include an item which asks about frustration level because an emotional aspect of the experience was not in the scope of the purpose of the study. As a result, the survey consisted of items that measure mental demand, temporal demand, overall performance, and effort.

In an attempt to capture more of users' perception of their own performance, I asked five questions about satisfaction on task performance on a 5-point scale: (1) perception of whether one found enough information, (2) satisfaction with the amount of time spent for the task, (3) satisfaction with the amount of information found, (4) satisfaction with the quality of information found and (5) satisfaction with the chosen strategy.

For locus of control, I adapted questions from the Multidimensional-Multiattributonal Causality Scale (MMCS) in Lefcourt et al (1979) that were designed to

assess the locus of control for achievement and were constructed from items representing each quadrant of the locus and stability of causal attribution model of B. Weiner et al (1971). The questions were adapted for the purposes of this paper to see one's beliefs about the causes of one's successes, which in this case means task completion. Several modifications made include deleting some items and modifying others to make them more appropriate for the context of web searching. I attempted to see if a user's perception of causal attribution for an observed outcome (i.e., their own search task performance) can vary depending on factors such as one's cognitive ability and task complexity, and if this shows any relationship with actual search outcome generated by a user. The survey was presented using Qualtrics software that allows participants to fill out the questionnaire online. The questionnaire took the participants about three minutes.

#### 4.6 Search behavior measures

Participants' search interaction data were recorded by a customized system used for the study. Measures taken from the log are in Table 3.

Table 3. Measures taken from search behaviors

duration	total amount of time spent during a session
queries	# of queries
query_length	# of unique terms per query
clicks	# of total clicks made on SERP
time_1 <sup>st</sup> click	average time taken for a participant between issuing a query and 1 <sup>st</sup> click on SERP (if any)

time_between_events	average time between all search activities (e.g., query issue, click, etc.)
---------------------	---

#### 4.7 Search outcome measures

In this study, participants were asked to complete a search task by doing their best to fill out a table. They were instructed to consider it as notes that they can refer to when it comes to a decision-making point. The instruction also indicated that it is totally up to them to do the following: a. what to put in a table and b. how much information they put in each cell. These participant-generated notes were considered as another dependent variable in addition to search behavior. I viewed it as an actual evidence of completeness of their performance as well as a visual representation of participants' search experience. I presented an example of search outcome in Table 4.

The idea of completeness can branch out into two operationalized concepts; breadth and depth. It is a commonly accepted idea that depth and breadth tradeoff exists in a variety of cases although both are necessary conditions for a quality outcome. In any circumstance where there is resource constraint that needs to be reconciled with the need, a tradeoff between breadth and depth occurs. Miller (1981) discussed the depth and breadth tradeoff in the context of human engineering evaluation, emphasizing that system designers should consider the optimal balance between depth (i.e., the number of menu levels) and breadth (i.e., the number of choices per menu). Pfeiffer (2012) considered depth and breadth of search as primary variables for consumer's decision-making behavior. Additionally, in their work of developing a model of data quality, Wang and Strong (1996, p. 32) defined completeness as "the extent to which data are of sufficient

breadth, depth, and scope of information contained in the data”. I found it as a good framework to incorporate in interpreting the search outcome formed in metrics table.

In this study, breadth was referred to as coverage which can be simply defined as the coverage of task facets in final notes. It was measured by the number of cells sufficiently filled with the relevant information over the total number of cells. The sufficiency threshold was set at 1, which means a cell should contain at least one piece of information (i.e., a discrete unit of information of a participant’s choice; phrases separated by comma or dash or bullet; sentences; links, etc.) to be counted as covered. Depth was measured in the following manner. First, I counted the number of pieces of information saved in an outcome table for each item across different dimensions then averaged those values. The final value indicates how deep a participant went to navigate and gather information for a specific item. The process of computing the depth also yielded an interesting insight into information cascade in distribution which shows the degree of unevenness or imbalance in the distribution of information in a table. All these measures are sorted and denoted in Table 4 below.

Table 4. Example of search outcome (Notes by participant 9 for simple version of task 1)

	Pug	Bichon Frise
ability to be left alone	<ul style="list-style-type: none"> <li>- Pugs are often called “shadows,” staying close to their owners’ sides at all times</li> <li>- Nap frequently but almost always ready to play</li> <li>- “Pugs do best in homes where they receive plenty of attention and are treated like members of the family, and in turn, they offer heaps of devotion and affection.”</li> </ul>	<ul style="list-style-type: none"> <li>- Said to be difficult to train in a house</li> <li>- “Your bichon will want to be near you all the time, so that is another positive if you like a dog on or near your lap”</li> <li>- Advised to not leave your Bichon alone for long periods of time; need to be kept in a crate</li> </ul>

need for outdoor activity	<ul style="list-style-type: none"> <li>- Pug owners would say that pugs are the “ideal house dog,” able to remain happy in the city or the country</li> <li>- Pugs are full of energy and personality</li> <li>- DO NOT require a lot of physical activity due to their small stature and lazy nature; too much activity might cause wheezing</li> <li>- “Pugs make great companions for those who live in apartments or homes without large backyards”</li> </ul>	- Requires a daily walk or two
suggestion	<p>Although the opinions differ among owners of each of the dogs, I would recommend getting a Pug rather than a Bichon Frise, as the Bichon seems as if it will have more of a difficulty staying home alone all day in the cage without a little exercise here and there. The pug does not require as much physical activity and therefore a backyard space. While the pug does require a lot of attention, I would say all dogs do, and you’ll be able to give that to the pug when you come home from work.</p>	

For instance, in this case, the coverage of the search outcome generated by participant 9 is 1 because all the cells were sufficiently covered. And, the depth will be 5.5  $((7+4)/2)$ . Lastly, the degree of imbalance will be 1.27  $(7/5.5)$ . This set of values can provide a good sense of how a participant assigned their effort in terms of generating search outcome (i.e., filling out a table), which is a primary part of task completion. Apart from the coverage, depth, and degree of imbalance, I also examined other measures from search outcome as indicated below in Table 5.

Table 5. Measures taken from search outcome

coverage	# of cells covered
depth	average # of pieces of information saved per item
degree of imbalance	maximum # of pieces of information saved for an item/ average # of pieces of information saved per item

notes_length	# of unique terms in table
justification_length	# of unique terms in suggestion section

## 5. Results

### 5.1 Working memory

Descriptive statistics of working memory span score of the sample (N=24) is provided in Table 6 and individual scores is presented in Figure 3. As one can see, participants' operation span scores covered a wide enough range and were nicely distributed within the range. For analysis, I divided participants into low- and high-groups using a median split. The operation scores were divided at 36 (low=0-36, high=37-60). This binary classification may result in loss of information, but given my low number of participants, I felt this would allow more reliable analysis.

Table 6. Descriptive statistics for operation span score

Descriptive statistics	possible range	Mean (SD)	Median	Min, Max
Operation span for a sample (N=24)	0-60	34.21 (13.15)	36	12, 60

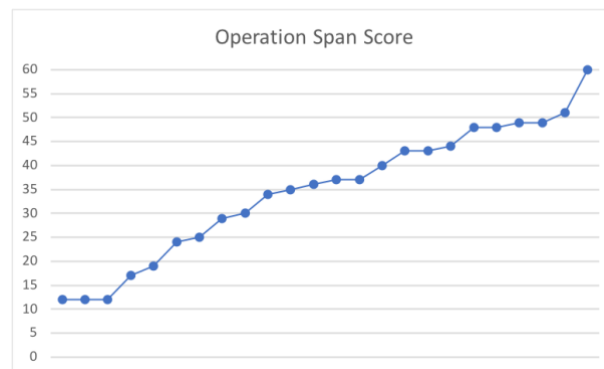


Figure 3. Distribution of operation span scores for participants (N=24)



## 5.2 Effects on users' perception

The alpha level for all analyses was set at .05. The central analysis for this study was a 2 (working memory: high and low) X 2 (level of task complexity: simple and complex) mixed-factor ANOVA. The main effects of task complexity and working memory capacity, as well as the interactions between those two factors, were examined, using dependent variables of measures of search behavior and self-reported responses from a post-task questionnaire. For the exploratory purpose of the study, I considered participants' responses to each TLX item.

My first research question (RQ1) investigates whether task complexity affects users' perception of their own performance while performing tasks of varying levels of complexity. I conducted one-way ANOVA analysis for each of measures I used in the post-task survey. Table 7 shows the means and standard deviations of all items, and the results of the ANOVA. Significant main effects were detected for temporal demand and satisfaction for the time spent on a task. None of the other measures showed a significantly strong relationship with task complexity. Participants reported greater temporal demand and lower level of satisfaction with time spent on a task during a complex task.

Table 7. Mean (Standard deviation) responses to the post-task survey items according to task complexity, F statistics and p-value

		Simple task	Complex task	F (p<0.05*)
TLX	Mental demand	2.58 (1.412)	2.88 (1.227)	F(1, 46)=.584, p=.449
	Temporal demand	2.54 (1.414)	3.38 (1.135)	F(1, 46)=5.072, p=.029*

	Performance	3.96 (.806)	3.63 (.770)	F(1, 46)=2.146, p=.150
	Effort	3.04 (1.268)	3.29 (1.197)	F(1, 46)=.493, p=.486
Satisfaction	Enough information	3.96 (1.042)	3.46 (1.250)	F(1, 46)=2.265, p=.139
	Time spent on the task	3.92 (1.100)	3.13 (1.154)	F(1, 46)=5.918, p=.019*
	Amount of information found	3.83 (1.167)	3.42 (1.176)	F(1, 46)=1.517, p=.224
	Quality of information found	3.79 (1.179)	3.50 (1.180)	F(1, 46)=.734, p=.396
	Strategies taken	4.17 (.761)	3.83 (.816)	F(1, 46)=2.140, p=.150
Locus of control	Internal_ability	4.92 (1.886)	4.83 (1.880)	F(1, 46)=.023, p=.879
	Internal_effort	5.46 (1.587)	5.33 (1.926)	F(1, 46)=.060, p=.807
	External_task	5.29 (1.334)	5.50 (1.319)	F(1, 46)=.296, p=.589
	External_system	4.54 (1.474)	4.21 (1.444)	F(1, 46)=.626, p=.433

My second RQ (RQ2) investigates whether working memory influences users' perception toward their own experience during a search task. No main effect of working memory capacity has been found from participants' responses to the post-task questionnaire. The F-statistics are not reported here to conserve space and were all smaller than 1.5. However, temporal demand of workload showed a relationship leaning toward significance [F(1, 46)=2.353, p=.132] and the ability dimension of internal locus of control almost attained significance [F(1, 46)=3.650, p=.062]. The mean values with 95% confidence interval for two variables are displayed in Figure 4.

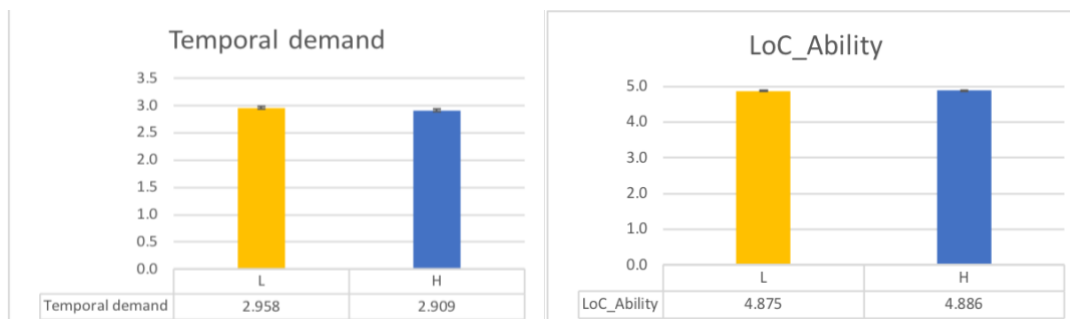


Figure 4. Trend analysis of working memory capacity for users' perceptions

To address RQ3, I examined the joint effect of task complexity and working memory on the users' perception. A two-way ANOVA yielded significant interaction effects for temporal demand [ $F(2, 45)=3.907, p=.027$ ] and marginal significance for satisfaction with the time spent on a task [ $F(2, 45)=2.967, p=.061$ ]. The effect of cognitive ability on perceived temporal demand and the level of satisfaction with time spent per task was stronger in the complex task condition. Figure 5 shows the means and 95% confidence intervals for those two post-task measures. No significant interaction effects were found for other measures (all  $F < 2$ ). The results show low- group participants experienced higher temporal demand than high- group participants, and the difference was greater in complex task condition. Also, participants reported a lower level of satisfaction with the time they spent on complex task than simple task, and the difference was more pronounced in low- group participants.

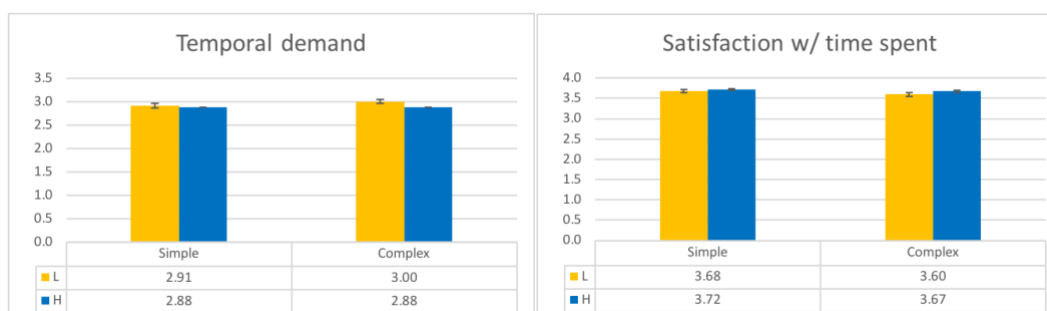


Figure 5. Interaction effects of task complexity and working memory capacity found on users' perception

### 5.3 Effects on users' search behavior

The same statistical analysis was conducted to see the effect of working memory and task complexity on each measure of search interactions (i.e., queries, a length of query, clicks, time spent until 1<sup>st</sup> click, session length, and time between events). For my first research question (RQ1), I investigated the main effect of task complexity on users' search behavior and no significant relationship was found for any of search interaction measures (all  $F < 1$ ) except a marginal significance found for queries [ $F(1, 46)=2.595$ ,  $p=.114$ ].

My second RQ (RQ2) examines the effect of working memory capacity on users' search behavior. Table 8 shows the means and standard deviations of all items, and the results of the ANOVA. Significant main effects were detected for queries, clicks, time taken until the first click, and time between events (i.e., query issues, clicks, etc.). To better understand the trends for the main effect of working memory capacity, Figure 6 is attached for the measures that returned significant relationship. Overall results strongly suggest that working memory capacity has an impact on users' search behavior. High-group participants exerted more effort than low- group participants by issuing more queries and more clicks, and low- participants took longer time until the first click and between search activities throughout the whole session. This trend of high- group participants performing activities but spending less time per activity suggests the idea of users with higher working memory performing a search task with greater efficiency. I

should also note that the effect sizes are very small; It is possible that this could be due to the use of median split in the analysis.

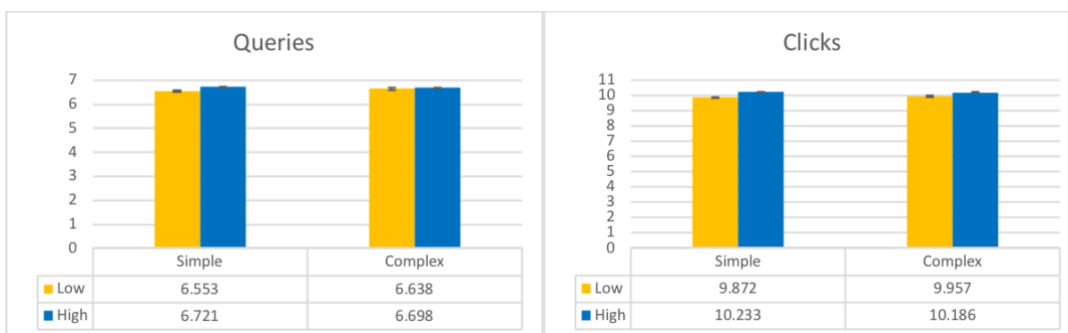
Table 8. Mean (Standard deviation) value for search interaction measures according to working memory capacity, F statistics and p-value

	Low	High	F (p<0.05*)
duration	689.667 (373.366)	694.273 (362.202)	F(1, 46)=.026, p=.874
queries	6.521 (4.100)	6.773 (4.153)	F(1, 46)=6.421, p=.015*
query_length	4.277 (1.656)	4.231 (1.629)	F(1, 46)=.578, p=.451
clicks	9.771 (5.459)	10.273 (5.389)	F(1, 46)=7.133, p=.010*
time_1 <sup>st</sup> click	9.785 (8.892)	9.632 (9.115)	F(1, 46)=5.165, p=.028*
time_between_events	47.577 (26.284)	45.619 (23.638)	F(1, 46)=6.421, p=.015*



Figure 6. Main effects of working memory on search behavior (i.e., queries, clicks, time until 1<sup>st</sup> click, and time between search activities)

For my third RQ (RQ3), I conducted a two-way ANOVA to understand if there is an interaction between the working memory capacity and task complexity for search behavior. Significant interaction effects were detected on queries [ $F(2, 45)=3.635$ ,  $p=.034$ ] and clicks [ $F(2, 45)=4.187$ ,  $p=.022$ ]. Also, marginal significance was found for time between events [ $F(2, 45)=3.141$ ,  $p=.053$ ]. Figure 7 shows the means and 95% confidence intervals for queries, clicks and time between events. No significant interaction effects were found for other measures (all  $F < 3$ ). Trend analysis show that main effect of working memory was greater in simple task condition for queries and clicks measures. However, in terms of time between search activities, the variance caused by working memory capacity was much greater under complex task condition than simple task condition. Low- group participants took much longer time between search activities than high- group participants especially when performing a complex task. Here also, effect sizes were found to be relatively small. Again, it is possible this may be attributable to the use of median split which can potentially cause a loss of effect size and power (Husser, 2017).



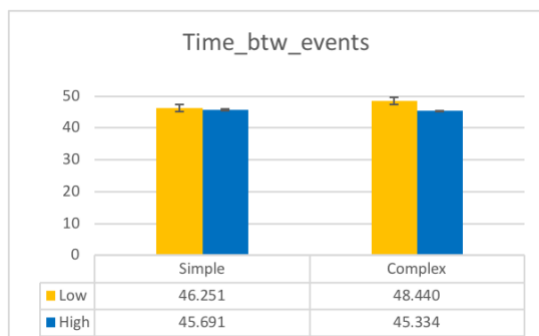


Figure 7. Interaction effects of task complexity and working memory capacity on search behavior (i.e., queries, clicks, and time between search activities)

## 6. Discussion

The findings from this study provided some insights about the relationship between working memory, task complexity and search behavior as well as users' perceptions about their experience. Task complexity had a significant main effect on users' perceived temporal workload (one dimension of workload considered in the NASA-TLX) as well as perceived level of satisfaction with time spent on a task. Yet, it didn't show any significant relationship with search behavior measures, which was not exactly in line with the findings of similar studies conducted (Brennan, Kelly & Arguello, 2014; Capra, Arguello, Crescenzi & Vardell, 2015). It is possible that users' perception was affected by confounding factors such as task topic and their prior knowledge about it, hindering the direct effect of manipulated factor (i.e., task complexity). In future studies, I intend to administer manipulation check to ensure that participants experienced the planned contrast between the two conditions. But, even then, the analysis revealed a significant effect of task complexity on users' time perception.

Working memory capacity had a significant main effect on several search behavior measures. High- group participants exhibited more search activities (i.e., queries and clicks) at a faster pace (i.e., shorter time until 1<sup>st</sup> click and shorter lag times between activities). This result was expected given that working memory capacity correlates comparatively highly with information processing speed (Kyllonen & Christal, 1990). It is worth noting that participants were instructed that they could stop searching any time when they thought they had enough information to complete the task; Yet, there was no



difference found in task duration between high- and low- group. High- participants carried out as much search activities during simple tasks as they did with complex tasks, while low- participants did less with less complex tasks. This suggests that users may have a different view or expectation in task completion according to their working memory capacity. In other words, task completion can be interpreted differently depending on users, rather than the number of requirements determine the amount of effort required for task completion. This may also suggest a correlation between working memory capacity and other constructs such as a need for cognition (Wu & Kelly, 2015).

With regard to locus of control measure, I assumed that low- group participants may experience more difficulties in performing complex task, possibly produce incomplete or partial search outcome and attribute the outcome to external cause such as task difficulty. On the other hand, I hypothesized that high- group participants would encounter fewer difficulties when seeking information, more likely complete a task to one's satisfaction and attribute the outcome (task success) to internal cause such as one's search ability or effort. However, I couldn't find any significant relationship between working memory and locus of control scale except that there was a marginal significance detected for the ability dimension of locus of control measures. General trend showed that high- group participants tended to attribute the outcome to one's search ability more than low- group participants.

One possible explanation for why locus of control measure did not show a significant relationship with either task complexity or cognitive ability is that because I did not include a half of the subscale. For the sake of simplicity, I only utilized items for success outcomes and not failure outcomes which can correspond to partial completion or

incomplete performance in the current study. The subscales of MMCS (Multidimensional-multiattributitional causality scales) contrasts the responses to success and failure and is designed to measure the attribution (causal belief) of both success and failure with respect to internal or external goal specific locus of control.

Lefcourt (1981) once argued that if one wished to create a locus of control index that would allow for precise predictions of human behavior, each of the subscales would be of vital concern. He also pointed out that positive and negative outcomes are often construed differently by the same persons; hence the responses to success, or positive outcomes, and failure, or negative outcomes should be assessed separately if only to make possible an exploration of the effects of such differential responding to events. Despite the concerns, I didn't strictly follow the given instructions, which can be acknowledged as a limitation of the study. I was under the assumption that all the participants will one way or another complete a whole task in a given time, which turned out not necessarily the case for everyone. So, asking if the main ingredient of task completion was one's search ability was not a proper question to ask to those who couldn't quite complete a task to their satisfaction. This will be reflected in my future studies.

Notwithstanding the limitation, the study shows interesting patterns of interaction dynamics between task complexity and working memory capacity on users' search behavior and perception of their performance. Working memory capacity had a stronger effect in complex task condition and this was primarily caused by low- group participants showing behavioral change. Additionally, even though the relationship was only marginally significant, low- group participants reported greater temporal demand than

high- group participants with both simple and complex tasks and they experienced even more temporal demands in complex task; whereas, high- group participants showed no difference in perceived temporal demand according to task complexity. Altogether, it suggests that task complexity may have a bigger impact on people with lower working memory capacity by imposing a greater challenge for them to cope with. However, caution must be applied in assessing the effect due to the small effect size. It could be partly because I used the median splits in data analysis, which could result in increases in Type II errors through loss of contrast (McClelland, Lynch, Irwin, Spiller & Fitzsimons, 2015).

I did not include analysis of participants' search outcome (i.e., notes and justification) for tasks in this paper. I plan to analyze the responses in future work. It will be interesting to see if there is any gap between participants' perception of their own performance and actual output they generated as they were performing tasks.

## 7. Conclusion

In this study, I examined the ways in which users' working memory capacity impacts their search behaviors and perceptions while conducting search tasks of varying complexity.

My work makes the following contributions to the IIR research community. First, it confirmed an effect of task complexity on perceived mental load (although it returned significance with only one subscale; temporal demand). Second, it also yielded supporting evidence of cognitive ability playing a key role in users' web searching. Third, it provided findings about the general adult population with good variability in working memory. Fourth, it used operation task for measuring one's complex working memory. To my knowledge, there has been no studies in the field that investigated working memory capacity using OSPAN which is designed to capture simultaneous storage and processing in conditions of interference that has high overlap with real-world information seeking situations. Fifth, it adopted the locus-of-control scale in task-based information retrieval setting as an attempt to understand users' attributional style for the causality of search task performance and outcome.

My findings reveal patterns of users with lower working memory experiencing greater demand, exerting fewer activities at a slower pace, and reporting a lower level of satisfaction. And, the impact of this cognitive ability was more prominent when users working on a more complex task. Although a significant relationship has been found, the analysis used here (i.e., median split) showed relatively small effect size difference, thus

more research is needed to examine the impacts more closely. Considering the usage of web search engines should not be restricted only to simple tasks but also be able to address users' complex need, the finding can call attention to the need for supplementary tools that could help users mitigate the difficulties they may experience during the complex search. Furthermore, since our cognitive abilities can show a decline with advanced age, design of search interfaces that enable customizing the interactions can be more salient in the future with an increase in life expectancy.

## **8. Acknowledgements**

This project would not have been possible without the help and support of Robert Capra and Jaime Arguello.

## 9. Reference

J. Arguello, W.-C. Wu, D. Kelly, and A. Edwards. Task complexity, vertical display and user interaction in aggregated search. In SIGIR, pages 435–444. ACM, 2012.

J. Arguello, [Predicting Search Task Difficulty](#). In Proceedings of the 36th European Conference in Information Retrieval (ECIR'14), 2014.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes. *Psychology of Learning and Motivation - Advances in Research and Theory*, 2(C), 89–195. [https://doi.org/10.1016/S0079-7421\(08\)60422-3](https://doi.org/10.1016/S0079-7421(08)60422-3)

Brennan, K., Kelly, D., & Arguello, J. (2014). The Effect of Cognitive Abilities on Information Search for Tasks of Varying Levels of Complexity. In Proceedings of the 5th Information Interaction in Context Symposium (pp. 165–174). New York, NY, USA: ACM. <https://doi.org/10.1145/2637002.2637022>

Campbell, D. J., 1988. Task Complexity: A Review and Analysis. *Acad Manage Rev*, 13, 1, 40-52.

Walhout, J., Oomen, P., Jarodzka, H., & Brand-Gruwel, S. (2017). Effects of Task Complexity on Online Search Behavior of Adolescents. *J. Assoc. Inf. Sci. Technol.*, 68(6), 1449–1461. <https://doi.org/10.1002/asi.23782>

Capra, R., Arguello, J., Crescenzi, A., & Vardell, E. (2015). Differences in the Use of Search Assistance for Tasks of Varying Complexity. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 23–32). New York, NY, USA: ACM. <https://doi.org/10.1145/2766462.2767741>

Case, R., Kurland, M. D., & Goldberg, J. (1982). Operational efficiency and growth of short-term memory span. *Journal of Experimental Child Psychology*, 33, 386-404.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, 30, 163 – 183.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769–786. <https://doi.org/10.3758/BF03196772>

- De Waard D. (1996). The Measurement of Drivers' Mental Workload. Ph.D. thesis, Traffic Research Center, Groningen University, Groningen.
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Dweck, C. S., & Reppucci, N. D. (1973). Learned helplessness and reinforcement responsibility in children. *Journal of Personality and Social Psychology*, 25, 109-116.
- Dweck, C. S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. *Developmental Psychology*, 14(3), 268–276. <https://doi.org/10.1037/0012-1649.14.3.268>
- Egan, D. E., Bowers, C., & Gomez, L. M. (1982). Learner characteristics that predict success in using a text-editor tutorial. *Proceedings of the Conference on Human Factors in Computing Systems*, Gaithersburg, Maryland.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, 128, 309-331.
- Fisk, J. E., & Warr, P. (1996). Age and working memory: the role of perceptual speed, the central executive, and the phonological loop. *Psychology and Aging*, 11(2), 316–323.
- Francis, G., Neath, I., & Mackewn, A. (n.d.). *Student Manual for CogLab™*, 122.
- Gao, Q. (2011). An Empirical Study of Tagging for Personal Information Organization: Performance, Workload, Memory, and Consistency. *International Journal of Human-Computer Interaction*, 27(9), 821-863.
- Gomez, L. M., Egan, D. E., Wheeler, E. A., Sharma, D. K., & Gruchacz, A. M. (1983). How interface design determines who has difficulty learning to use a text editor. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Boston, MA.
- Gomez, L. M., Egan, D. E., & Bowers, C. (1986). Learning to use a text editor: Some learner characteristics that predict success. *Human Computer Interaction*, 2(1), 1-23.
- Gwizdka, J. (2009). What a difference a tag cloud makes: Effects of tasks and cognitive abilities on search results interface use. *Information Research*, 14(4).
- Gwizdka, J. (2013). Effects of working memory capacity on users' search effort. In *Proceedings of the International Conference on Multimedia, Interaction, Design and Innovation* (p. 11). ACM.



Gwizdka, J. (2017). I Can and So I Search More: Effects of Memory Span On Search Behavior. In Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval (pp. 341–344). ACM.

Haapalainen, E., Kim, S. J., Forlizzi, J. F., & Dey, A. K. (2010). Psycho-physiological measures for assessing cognitive load. Proceedings of the ACM International Conference on Ubiquitous Computing, Copenhagen, Denmark.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: NH Press.

Hüsser, A. P. (2017). Median Split. In *The International Encyclopedia of Communication Research Methods* (pp. 1–2). American Cancer Society.  
<https://doi.org/10.1002/9781118901731.iecrm0148>

B. J. Jansen, D. Booth, and B. Smith. Using the taxonomy of cognitive learning to model online searching. *Inf. Process. Manage.*, 45(6):643–663, Nov. 2009.

Jex, H. R. (1988). Measuring mental workload: Problems, progress, and promises. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (Vol. 52, pp. 5-39). Amsterdam, The Netherlands: Elsevier Science Publishers B.V.

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133, 189-217.

Kelton, A. S., & Pennington, R. R. (2012). Internet financial reporting: The effects of information presentation format and content differences on investor decision making. *Computers in Human Behavior*, 28(4), 1178–1185.

Klein, K., & Fiss, W. H. (1999). The reliability and stability of the Turner and Engle working memory task. *Behavior Research Methods, Instruments, & Computers*, 31, 429-432.

Kuhlthau, C.C. (1919). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42, 361-371.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic Performance, Career Potential, Creativity, and Job Performance: Can One Construct Predict Them All? *Journal of Personality and Social Psychology*, 86(1), 148-161

Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences*, 2, 129–160.

Lefcourt, H. M., von Baeyer, C. L., Ware, E. E., & Cox, D. J. (1979). The multidimensional-multiattributitional causality scale: The development of a goal specific locus of control scale. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, 11(4), 286–304. <https://doi.org/10.1037/h0081598>

Lefcourt, H. M. (1981). The construction and development of the multidimensional-multiattributitional causality scales. In Lefcourt, H. M. (Eds.), *Research with the locus of control* (pp. 245-275). New York, USA: Academic Press.

MacFarlane, A., Albrair, A., Marshall, C., & Buchanan, G. (2012). Phonological working memory impacts on information searching: An investigation of dyslexia. *Proceedings of the 4th Information Interaction in Context Symposium*, Amsterdam, The Netherlands.

Marchionini, G. (1995). *Information Seeking in Electronic Environments*. New York, NY, USA: Cambridge University Press.

McClelland, G. H., Lynch, J. G., Irwin, J. R., Spiller, S. A., & Fitzsimons, G. J. (2015). Median splits, Type II errors, and false-positive consumer psychology: Don't fight the power. *Journal of Consumer Psychology*, 25(4), 679–689. <https://doi.org/10.1016/j.jcps.2015.05.006>

Miller, G. A. (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information". *Psychological Review*. 63 (2): 81–97. doi:10.1037/h0043158. PMID 13310704

Miller, D. P. (1981). The Depth/Breadth Tradeoff in Hierarchical Computer Menus. *Proceedings of the Human Factors Society Annual Meeting*, 25(1), 296–300. <https://doi.org/10.1177/107118138102500179>

Pfeiffer, J. (2012). Fundamentals on Decision-Making Behavior. In *Interactive Decision Aids in E-Commerce* (pp. 15–45). Physica-Verlag HD. [https://doi.org/10.1007/978-3-7908-2769-9\\_2](https://doi.org/10.1007/978-3-7908-2769-9_2)

Rohde T. E., Thompson L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence* 35 83–92. 10.1016/j.intell.2006.05.004

Rotter, J. B. (1954). *Social learning and clinical psychology*. New York: Prentice-Hall. <http://dx.doi.org/10.1037/10788-000>

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80(1, Whole No. 609).

- Ruffing, S., Wach, F.-S., Spinath, F. M., Brünken, R., & Karbach, J. (2015). Learning strategies and general cognitive ability as predictors of gender- specific academic achievement. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01238>
- Salthouse, T. A. (1993b). Speed mediation of adult age differences in cognition. *Developmental Psychology*, 29, 7 22-7 38
- Schmitt, N. (2014). Personality and Cognitive Ability as Predictors of Effective Performance at Work. *Annual Review of Organizational Psychology and Organizational Behavior*, 1(1), 45–65. <https://doi.org/10.1146/annurev-orgpsych-031413-091255>
- Silva, F. P. da. (2014). Mental Workload, Task Demand and Driving Performance: What Relation? *Procedia - Social and Behavioral Sciences*, 162, 310–319. <https://doi.org/10.1016/j.sbspro.2014.12.212>
- Speier, C., & Morris, M. G. (2003). The influence of query interface design on decision-making performance. *MIS Quarterly*, 27(3), 397-423.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, 28, 127-154.
- Turpin, L., Kelly, D., & Arguello, J. (2016). To Blend or Not to Blend?: Perceptual Speed, Visual Memory and Aggregated Search. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 1021–1024). New York, NY, USA: ACM. <https://doi.org/10.1145/2911451.2914739>
- Wang RY, Strong DM. Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems*. 1996;12:5–34.
- Weiner, B. (1972). Attribution Theory, Achievement Motivation, and the Educational Process. *Review of Educational Research*, 42(2), 203–215. <https://doi.org/10.3102/00346543042002203>
- Wilson, T. D. (2000). Human Information Behavior. *Informing Science The International Journal of an Emerging Transdiscipline*, 3(2), 49–56.
- Wood, R.E., 1986. Task complexity: Definition of the construct. *Organ Behav Hum Dec*, 37, 1, 60-82.
- Wu, W.-C., & Kelly, D. (n.d.). Online search stopping behaviors: An investigation of query abandonment and task stopping. *Proceedings of the American Society for Information Science and Technology*, 51(1), 1–10. <https://doi.org/10.1002/meet.2014.14505101030>

## Appendix 1. Consent form

University of North Carolina at Chapel Hill  
Information about a Research Study  
Adult Participants

Consent Form Version Date: 01/18/2018

IRB Study # 17-3304

Title of Study: The effect of cognitive style on user's search behavior and experience during interacting with tasks with different level of complexity

Principal Investigator: Bogeum Choi

Principal Investigator Department: School of Information and Library Science

Principal Investigator Phone number: [REDACTED]

Principal Investigator Email Address: choiboge@email.unc.edu

Funding Source and/or Sponsor: N/A

---

What are some general things you should know about research studies?

You are being asked to take part in a research study. To join the study is voluntary. You may choose not to participate, or you may withdraw your consent to be in the study, for any reason, without penalty.

Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about taking part in this research study. You should ask the researchers named above, any questions you have about this study at any time.

What is the purpose of this study?

The purpose of this study is to investigate how individual differences can affect search behaviors and experience while the one is interacting with search tasks of different levels of complexity.

You are being asked to participate in this study because you probably routinely use search engines such as Google to find information.

Are there any reasons you should not be in this study?

- You should not be in this study if you have never used a search engine such as Google.
- You should not be in this study if you are younger than 18 years old.

- You must be fluent in English.
- You must work at UNC

How many people will take part in this study?  
A total of approximately 24 people will take part in this study.

How long will your part in this study last?  
The whole study session will take approximately an hour.

What will happen if you take part in the study?  
First, you will perform two search tasks that will require you to search information that help you complete given tasks, using the given system. Then you will be asked to complete a post-task survey after each task. Then you will complete a memory task at the end.

What are the possible benefits from being in this study?  
Research is designed to benefit society by gaining new knowledge. You will not benefit personally from being in this research study.

What are the possible risks or discomforts involved from being in this study?  
You may experience embarrassment if you are unable to find useful information for an assigned search task. To mitigate this risk, we are using search tasks that have been used in previous studies of search behavior and represent typical search tasks that people complete on a daily basis.

How will information about you be protected?  
Data generated from the interaction between you and the system (e.g., clicks, queries) will be recorded using a web browser plug-in. The data collected will be stored on password protected computers of the research team, Qualtrics servers and on UNC OneDrive. We will use a label to identify the data, not your name.

Will you receive anything for being in this study?  
You will receive \$15 in an exchange of participation.

What if you have questions about this study?  
You have the right to ask, and have answered, any questions you may have about this research. If you have questions about the study, complaints, concerns, or if a research-related injury occurs, you should contact the researchers listed on the first page of this form.

What if you have questions about your rights as a research participant?  
All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject, or if you would like to obtain information or offer input, you may contact the Institutional Review Board at 919-966-3113 or by email to [IRB\\_subjects@unc.edu](mailto:IRB_subjects@unc.edu).

Additionally, you may contact the investigators working on this study: Bogeum Choi  
([choiboge@email.unc.edu](mailto:choiboge@email.unc.edu))

Participant's Agreement:

I have read the information provided above. I have asked all the questions I have at this time. I voluntarily agree to continue to participate in this research study.

---

Signature of Research Participant

Date

---

Printed Name of Research Participant

---

Signature of Research Team Member Obtaining Consent

Date

---

Printed Name of Research Team Member Obtaining Consent

## Appendix 2. General instructions page

### Task instruction

- Read the task description
- Look for the information to complete the task using the given system
- Fill out the table on the word document using the information you found useful
- You will have to make a decision at the end based on the information you found

### Example

Your friend recently moved to the town of Chapel Hill and he asked you about a grocery store near him to shop for food. Now you are trying to help him find make a decision.

How do a) Food lion and b) Whole food market differ as a choice for your friend in terms of price range?

	Food lion	Whole food market
price range	Notes (you can put whatever you will find useful) - link of a webpage - text from a webpage - your own notes	
suggestion to your friend	I would suggest... because... (brief justification based on the information you found above)	

- It is up to you to decide what to put in the table
- Your goal is to find as much information as you think is needed to address each cell in the table; some cells may require multiple pieces of information
- You will have up to 20 mins to complete the task and I will give you 2-minute warning

### Appendix 3. Search tasks

Topic	Level of complexity	Description
Dog adoption	simple (2 items X 2 dimensions)	A friend of yours has recently decided to get a dog for companionship. Your friend works during the day and lives in an apartment complex. Now you are trying to help your friend make a decision. How do a) Pug and b) Bichon Frise dog breeds differ as a choice for your friend in terms of a) its ability to be left alone and b) need for outdoor activity?
	complex (4 items X 4 dimensions)	A friend of yours has recently decided to get a dog for companionship. Your friend works during the day and lives in an apartment complex. Now you are trying to help your friend make a decision. How do a) Pug, b) Bichon Frise, c) Beagles and d) Golden Retriever dog breeds differ as a choice for your friend in terms of a) its ability to be left alone, b) need for outdoor activity, c) how well they do in small space and d) temperament?
Ways to quit smoking	simple (2 items X 2 dimensions)	A friend of yours has recently decided to quit smoking and asked for your help in choosing a method. How do a) nicotine replacement therapy (NRT) and b) medication work differently in terms of a) success rate and b) cost?
	complex (4 items X 4 dimensions)	A friend of yours has recently decided to quit smoking and asked for your help in choosing a method. How do a) nicotine replacement therapy (NRT), b) medication, c) behavioral therapy and d) e-cigarette work differently in terms of a) average treatment length, b) cost, c) any side effect and d) success rate?



### Appendix 4. Post-task survey

#### ID

Post-task survey

Enter participant ID:

#### Workload

How mentally demanding was the task?

Very low     Very high

How hurried or rushed was the pace of the task?

Very low     Very high

How successful were you in accomplishing what you were asked to do?

Failure     Perfect

How hard did you have to work to accomplish your level of performance?

Very low     Very high

#### Satisfaction

I found enough information about the search task

Strongly disagree  Disagree  Neutral  Agree  Strongly agree

I am satisfied with the amount of time I spent for the task

Very dissatisfied	Dissatisfied	Neither satisfied nor dissatisfied	Satisfied	Very satisfied
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

I am satisfied with the amount of information found for the task

Very dissatisfied	Dissatisfied	Neither satisfied nor dissatisfied	Satisfied	Very satisfied
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

I am satisfied with the quality of information found for the task

Very dissatisfied	Dissatisfied	Neither satisfied nor dissatisfied	Satisfied	Very satisfied
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

I am satisfied with the strategies I took to find information about the task

Very dissatisfied	Dissatisfied	Neither satisfied nor dissatisfied	Satisfied	Very satisfied
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Locus of control**

The most important ingredient in this task completion was my search ability

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

I felt that my search ability reflected directly on my task success

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In this task, the useful information I found was the direct result of my search effort

Strongly disagree	Disagree	Neutral	Agree	Strongly agree
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

I was able to find useful information because I put enough effort into this task

Strongly disagree      Disagree      Neutral      Agree      Strongly agree  
                                                                                       

Some of the times that I found useful information, it was due to the easiness of the task

Strongly disagree      Disagree      Neutral      Agree      Strongly agree  
                                                                                       

Some of the useful information I found may simply reflect that it was easier to retrieve than most

Strongly disagree      Disagree      Neutral      Agree      Strongly agree  
                                                                                       

My success on this task depended on the search system's capabilities

Strongly disagree      Disagree      Neutral      Agree      Strongly agree  
                                                                                       

I felt that some of the useful information I found depended on chance factors, such as luck

Strongly disagree      Disagree      Neutral      Agree      Strongly agree