Mental States in Information Search Process

정보검색과정에서의 정신적 상태

나 경 식(Kyoungsik Na)*
최 원 태((Wontae Choi)**

〈목　차〉

Ⅰ. Introduction
Ⅱ. Related Studies
1. Cognitive load and mental states
2. Information search
Ⅲ. Research Questions
Ⅳ. Methods
1. Experimental design
2. Participants
3. Dependent measures
4. Procedure and search tasks
Ⅴ. Data Analysis and Discussion
1. Mental States
2. NASA-TLX Cognitive Load
3. Correlations between Each of the Mental State Indicators
Ⅵ. Conclusion and Implications
Ⅶ. Limitations and Future Direction

ABSTRACT

This study examined the differences in a searcher's mental states of a complex information search and retrieval task during information search process between the two groups: participants who were exposed to mental demand manipulation and those not exposed. Data from the experiments and questionnaires were analyzed. Based on qualitative approach and quantitative analyses, the results indicated that the participants exposed to mental demand required more thoughts; addressed negative emotions more often; reduced a searcher's efforts; and interrupted search performance than those not exposed. These results suggest that mental demand contributed to a searcher's perceived thought, emotion, effort, and performance, although these mental states differed in relative contribution of information search process. Significant differences were found between the two groups with respect to the component of mental demand, performance, and frustration of the NASA-TLX subjective cognitive load. These results have implications for search user interface design and information search systems among others.

Keywords: Mental states, Cognitive load, Mental demand, NASA-TLX

초 록

본 연구는 두 집단간에 복잡한 정보검색임무를 수행하는 동안 이용자의 정신상태의 차이점을 살펴보고자 하였다. 실험과 질문지를 통해서 데이터를 수집하고 분석하였고, 그 결과 정신적요구를 받은 참여자들이 안받은 참여자들보다 검색 중 생각을 더 많이 한 것으로 보이고, 부정적인 감정표현을 더 자주 한 것으로 보이고, 노력을 덜 한 것으로 보이며, 검색수행능력이 훼방을 받은 것으로 드러났다. NASA-TLX의 검사결과는 정신적요구, 수행능력, 좌절감에서 주요한 차이를 보이는 것으로 나타났다. 이러한 연구결과는 인터페이스 디자인이나 정보검색 시스템등의 개발에 도움을 줄 것으로 기대한다.

키워드: 정신적상태, 인지부하, 정신적요구, 정보검색

* 건국대학교 문헌정보학과 부교수(kn05d@kku.ac.kr) (제1저자)
** 건국대학교 문헌정보학과 교수(wtchoi@kku.ac.kr) (교신저자)
•논문접수: 2017년 8월 20일 •최초심사: 2017년 8월 29일 •게재확정: 2017년 9월 21일
•한국도서관·정보학회지 48(3), 281-302, 2017. [http://dx.doi.org/10.16981/kliss.48.201709.281]

Ⅰ. Introduction

People experience cognitive activities when they find information to solve problems interacting with information systems. Information search and retrieval is considered as a complex and also learning process, and thus demands a variety of human cognitive activities not only from mental efforts but also from their physical and behavioral activities (Gimino 2002; Annett 2002). Information search on the Web spends cognitive load because searchers could consume mental and physical activities, effort or performance throughout a course of search with a search system. Cognitive load is defined as "the total amount of mental efforts being used in the limited amount of working memory at a given time (Sweller 1988; 2010)." Cognitive load theory was developed by John Sweller (1988) indicating that there are three types of cognitive load: intrinsic, extrinsic, and germane cognitive load. In educational settings, it is preferable to reduce extrinsic cognitive load and increase germane cognitive load for an optimum learning outcome (Kirschner 2002).

It is widely accepted that a heavy cognitive load contributes to poor performance (Craik and Lockhart 1972; Pass et al. 2005; Pass et al 2003). Mental demand is often described as being one of the factors of cognitive load. It also refers to one of the six components in the NASA-TLX self-workload assessment tool developed by Staveland and Harts. People are able to cope with a certain level of mental demand within certain limits, but unable to deal with its overload when the level of mental demand is too high. It is argued that the effects of mental demand are diverse such as stress, emotional instability, psychological status, or mental and time pressures placed on an individual (Oosterwijk et al. 1996). Thus, errors might occur when these pressures place upon an individual who are unable to cope with it (Oosterwijk et al. 1996).

Today, with the development of related information technologies, information search on the Web is growing fast and being used to solve problems for their information need, and thus it has become a daily routine behavior. Past studies of information search and retrieval have extensively focused on performance of information systems because of the fact that a search was measured whether it could find relevant information at the time of a search within a desired time based on recall and precision. However, research on cognitive load in the area of information retrieval still lacks although there is a continuous interest regarding the effects of cognitive load that might influence searchers' information search experience (Gwizdka 2008; 2010; 2013; Zhang and Gwizdka 2014). Understanding of characteristics of a user is equally important when designing and developing information search systems for a user-friendly interaction.

Ⅱ. Related Studies

1. Cognitive Load and Mental States

Cognitive load is an important factor that should be considered in the design and development of information systems (Gwizdka 2010). Cognitive load differentiates three types: intrinsic, extrinsic, and germane cognitive load (Sweller 1988). In the context of information retrieval, intrinsic cognitive load depends on the levels of complexity of the given information search task. Extrinsic cognitive load depends on information presentation being processed in a searcher’s relevance judgment (Salomon 1984), whereas germane cognitive load results from a searcher’s engagement such as efforts to put in the process of information search. It is argued that cognitive load can be influenced by various factors that make a definitive measurement difficult (Pass et al. 2003). Bruken, Plass, and Leutner (2003) stated that cognitive load can be treated as “a theoretical construct, describing the internal processes of information processing that cannot be observed directly” (p. 55). According to them (2003), there are various methods to assess cognitive load that are currently available along with two dimensions, subjective vs. objective and direct vs. indirect (See Tab. 1). Among those methods, a dual-task approach and the NASA-TLX measurement are well known and often used to assess cognitive load.

The dual-task method has been used to induce extra mental memory load when a secondary task was added to a primary task. The NASA-TLX measurement was originally developed by Hart and Staveland as a perceived workload index and used to measure subjective cognitive load from six components, namely, mental demand, physical demand, temporal demand, effort, frustration, and performance (2017). One limitation of the NASA-TLX subjective tool is that it relied heavily on self-reported measures. To this end, more objective tools as a direct way for measuring mental workload such as eye-tracking method, brain activity measures, and physiological measures were used to compare the differences between subjective and objective measures (See Tab. 1).

Niculescu et al. (2010) studied the impacts of stress and cognitive load on the perceived interaction quality of a multimodal system indicating that both “factors had an influence on the way subjects perceived the interaction quality, whereas the cognitive load seems to have a higher impact” (p.453). Although cognitive load is typically a study area in education, psychology, or ergonomics, there has been a growing interest in the area of Information Retrieval (Gwizdka 2008; 2010).

〈Tab. 1〉 Methods for cognitive load measurement (Adopted from Bruken, Plass, and Leutner, 2003)

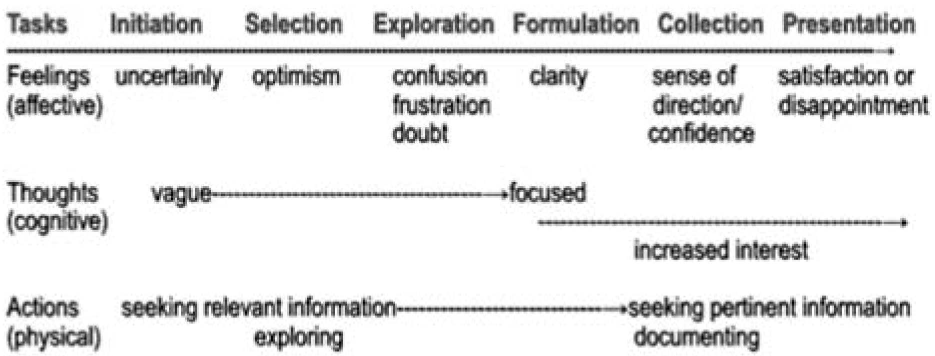
	Indirect	Direct
Subjective	Self-reported invested mental effort	Self-reported stress level
		Self-reported difficulty of materials
Objective	Dual-task performance	Brain activity measures (e.g., fMRI)
	Behavioral measures	Physiological measures
	Learning outcome measures	

Mental states are often studied in an area of cognitive psychology to examine one’s information processing that depends on and also corresponds to complex characteristics such as thinking, feeling, remembering, emotions, and so on (Salomon 1984). Clore and Huntsinger (2007) studied how emotions inform judgment and regulate thought indicating that affect and emotions are pervasive influences on human judgment and thought. They (2007) also stated that “positive affective information promotes and negative affective information inhibits the cognitive responses that are accessible or dominant in a particular situation” (p.397). Positive thinking may increase motivation through raising effort that will lead to good performance obtaining a good result.

Although researchers have examined for the physical correlates of these mental states in different disciplines (Howells et al. 2010), mental states research in the area of information retrieval is still lacking. In this study, we focus on how our different categories of mental states interact with a search system on what happens when people search for relevant information (Howells et al. 2010).

2. Information Search

Studies of Information search have long been researched focusing on many different areas such as search techniques, query behavior, search user interface, relevance judgment, and task complexity, among others (Bates 1989; Byström 2002; 2005; Kirschner et al. 2009; 2011; Vakkari and Pertti 2001; Vakkari 1999; Xie and Salvendy 2000). There are many other important factors that need to be addressed from the perspective of users along with a system side (Wilson 2006). Kuhlthau's model of information search process (ISP) was developed in the 1980s and the model has been used as a framework for understanding the information search experience of people various library and information science settings (Kuhlthau et. Al 2008). Kuhlthau and her colleagues (2008) examined the continued usefulness of Kuhlthau's information search process as a model of information behavior in new, technologically rich information environments. And they found that the information search process model remains still useful for explaining student's information behaviors.



<Fig. 1> Model of the Information Search Process (Adopted from Kuhlthau et. al, 2008)

User's search experience with a search engine is a good example as well as performance, task effectiveness, satisfaction, error rate, and utilities (Pu and Kumar 2004). Usability testing is another important method to assess how easy and useful user interfaces are to use (Nielson 2017). In this study, we focus on how mental demand affects information search process for a user interaction with information search systems.

Ⅲ. Research Questions

The overarching question is what happens when a searcher experiences mental demand with respect to a searcher's perceived thoughts, emotions, efforts and performance in the process of information search. In this paper, the reason we chose these four categories is that they may represent the core actions within general Information search and retrieval tasks interacting with the information search systems during information search process. These four categories were also derived from the elements of the domains of Kuhlthau's information search process (See Fig 1). Feelings, thoughts, and actions were adopted and in the three elements in Kuhlthau's ISP. Effort category was added in our study because mental effort ratings and secondary task performance have shown promise as measures of cognitive load in healthcare information system settings (Haji et. al 2015). It is said that mental state examination in psychology is an important part of the clinical assessment observing and describing one's psychological functioning at a given point in time under the domains of appearance, attitude, behavior, mood and affect, thought process and content, perception, cognition, insight and judgment (Trzepacz and Baker 1993).

To this end, we examined these four categories in association with mental demand whether or not there were any differences during information search between the two groups. The present study therefore seeks to answer the following research questions:

1) Are there any differences in searchers' mental states in between those exposed to mental demand and those not exposed?
2) Are there any differences in searchers' cognitive load in between those exposed to mental demand and those not exposed?
3) Are there any relationships among each category of thoughts, emotions, efforts, or performance?

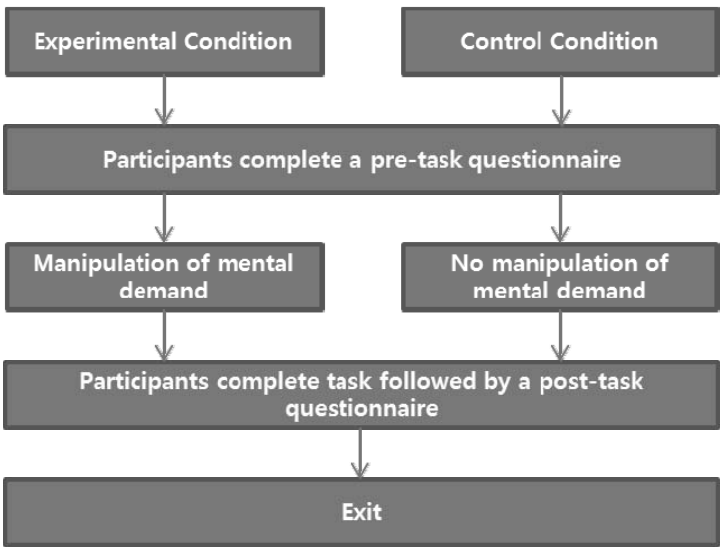
According to this first research question, for mental states, it is assumed that a searcher without mental demand manipulation generates a lower degree of thoughts, positive emotions, less efforts, and better performance in contrast to a searcher with mental demand manipulation due to the cognitive load. It is assumed that thoughts, emotion, effort, and performance could be hindered with extra mental demand manipulation, and positive emotion could influence to yield more efforts that might lead to a better performance. As for the second research question, for cognitive load, it is assumed that a searcher without mental demand manipulation show decreased cognitive load in total in comparison to a searcher with mental demand manipulation.

Ⅳ. Methods

1. Experimental Design

All participants took part in the experiment in one of the two conditions: experimental vs. control condition. In both conditions, they were asked to perform a search task and complete a pre- and a post-task questionnaire. We used think-aloud protocol by taking a dual-task method to examine the participants' mental states. A repeated measures design was utilized to investigate the effects of mental demand on participants' perceived thoughts, emotions, efforts, and performance during information search process. Mental demand was manipulated at one level involving memorize and recall test (MRT). MRT has been shown to cause mental demand in the study of cognitive load and query reformulation behavior.

To increase extra mental demand, a dual-task method was employed. In the dual-task of this study, the primary task was a search task, while the second task was MRT. The participants in the experimental group were asked to memorize an 8 digit numbers such as 94756294 and recall it at the completion of the search task. The control group remains the same except the mental demand manipulation.



〈Fig. 2〉 Conditions of Experiment

2. Participants

Thirty-six university students were recruited for this study. Mean±SD age was 22.58±1.50 years and there were 18 males and 18 females. Participants who drew even numbers were randomly assigned to the experimental group and those with odd numbers were assigned to the control group.

When asked about computer knowledge, 32 of the 36 (91.7%) participants reported their computer knowledge was more than average. When asked about web search knowledge, 33 of the 36 participants (36.4%) reported that their web search knowledge was more than average. When asked about web search success, 8 of the 36 participants (22.2%) reported it rarely successful, 17 of the 36 participants (47.2%) found it sometimes successful, and 11 of the 36 participants (30.6%) found it usually successful, respectively. Participants revealed no demographic differences in terms of age and search experiences between the two groups.

3. Dependent Measures

To collect participants' demographic information, a pre-task questionnaire was used. To quantify perceptions of workload, the National Aeronautics and Space Administration-Task Load Index (NASA-TLX) was modified and used in a post-task questionnaire. To collect measurements, we used two questionnaires and experiments. Four categories of searchers' perceived components, namely, thoughts, emotions, efforts, and performance, were examined in the experiments using think-aloud protocol and a dual-task method. All experiments were screen-captured using Camtasia program and later transcribed.

1) Mental state measurement

Mental states are multidimensional and complex phenomena, and thus difficult to directly examine. Therefore, we examined differences in mental states indirectly across four dimensions that were predetermined four set of categories in the process of the experiments: thoughts, emotions, efforts, and performance:

(1) Thoughts: Searchers' perceived thoughts are defined as activities that require human cognition and perceptions. Any activities that fall under this category such as thinking,

remembering, relevance judgment, searching, or retrieval were coded into searchers' thoughts category (Blumberg 2000; Oosterwijk et al. 2012; Peterson et al. 2009).

(2) Emotions: Considering searcher's perceived emotions, we identified participants' verbalized terms associated with searchers' emotions, which are subjective, conscious experience characterized by psychological expressions, biological reactions, and mental states such as confidence, frustration, stress, and confusion in this study (Bandura 1982; Seibert and Ellis 1991).

(3) Efforts: Searcher's perceived efforts are defined activities that use a series of actions requiring physical or mental energy. In regards to searcher's perceived efforts, we identified terms associated with time spent, queries, pages visit, or work hard.

(4) Performance: Searcher's perceived performance was characterized by a searcher's expressions or accomplishments of actions, competence, satisfaction, or efficacy in a given task. We identified two aspects of analysis, i.e., positive expressions (i.e., high level) and negative expressions (i.e., low level).

These four variables were later separated into subsections (see Tab. 2). These sub-sections were derived from Kuhlthau's information search process activities (1991). The frequency of these four subcategories of term occurrences was recorded and later coded and examined for data analysis. To measure these dependent variables, a concurrent think-aloud protocol was employed. During the think-aloud protocol, participants were encouraged to verbalize their thoughts, emotions, efforts, and performance as they processed a search task. During the whole search process, the participants were reminded to verbally state what they were thinking, how they were feeling, and how hard they would work and perform the search task. The verbal statements were recorded and then transcribed and encoded for later analysis (See Tab. 2).

〈Tab. 2〉 Coding scheme

	Mental States Category			
	Thoughts	Emotions	Efforts	Performance
Subsections (Indicators)	Thinking	Confidence	Time spent	performance (positive)
	Remembering	Frustration	Query made	performance (negative)
	Searching	Stress	Pages visit	satisfaction (positive)
	Relevance judgment	Confusion	Work hard	satisfaction (negative)

1) Cognitive load measurement

Data was also collected by a post-task questionnaire to collect cognitive load score using the

NASA-TLX subjective workload assessment. We therefore used both qualitative approach and quantitative method.

4. Procedure and Search Tasks

The present study was conducted in a controlled laboratory setting. An exploratory user study was employed including a pre-task questionnaire, experiments, and a post-task questionnaire in this order. At the beginning, participants were asked to sign the consent form and had an introductory session. After the introductory session, participants completed a pre-task questionnaire that collects their search experiences. And then they performed the first search task and at completion of the search task, they completed a post-task questionnaire in order to collect their workload of NASA-TLX. After 10 minute break to avoid mental fatigue, they performed the second search task followed by a post-task questionnaire. Search tasks were scenario-based and exploratory that might require searchers' cognitive activities.

Ⅴ. Data Analysis and Discussion

1. Mental States

An ANOVA was performed to measure the overall difference in the mean level of four categories of mental states and type of cognitive load between the groups. Regarding demographic characteristics of the participants, we analyzed group differences between the two groups in demographic variables and questionnaire subscale scores using ANOVA as well as Chi square test for categorical data. In order to test hypotheses, we performed repeated measures analysis of variance for differences in thoughts, emotions, efforts, and performance.

1) Searchers' Perceived Thoughts

The first research question looked at searcher's mental states during information search process. Qualitative analysis was conducted using a think-aloud protocol via the content analysis method by Garrison and others (Garrison et al. 2006). As discussed earlier, we identified themes associated with thoughts, which were mostly related to human cognition and perceptions. Certain comments may be explained about searchers' thoughts during information search process. For example:

"I am thinking keywords as to what I should find relevant information. It made me think continuously during information search over and over again. I am searching for this keyword again to find relevant information. There are unrelated results on my first search, it made me try second search to find more relevant information."

These excerpts show that participants are aware of a burden of more reactions about their cognition and perception. Think-aloud data also give insight in the way participants react to perceived thoughts. For instance, participants reported thoughts about query formulation, search process, or relevance judgment.

For searcher's perceived thoughts in quantitative analysis, an ANOVA was conducted to examine whether there were differences between the two groups. The results indicated that the participants with mental demand did express thoughts more often than those without mental demand in frequency of thoughts term occurrences based on the subsections of coding scheme (See Tab. 3). Among these four sub-sections, there were significant differences between the two groups in terms of thinking, remembering, and relevance judgment. This could imply that mental demand could require more thinking and remembering, and issue relevance judgment frequently,

〈Tab. 3〉 Summary of perceived thoughts

	Thinking											
	Task1						Task2					
	M	SD	min	max	F	p	M	SD	min	max	F	p
Ex	5.22	2.29	2	9	14.207	.001*	4.56	2.57	3	22	7.247	.011*
Co	2.50	2.04	1	8			2.56	1.82	1	9		
	Remembering											
	Task1						Task2					
	M	SD	min	max	F	p	M	SD	min	max	F	p
Ex	9.44	4.58	3	18	64.493	.000*	5.78	3.15	2	14	41.404	.000*
Co	0.56	1.04	0	4			.56	1.38	0	5		
	Relevance judgement											
	Task1						Task2					
	M	SD	min	max	F	p	M	SD	min	max	F	p
Ex	14.00	6.09	9	34	10.972	.002*	11.39	5.69	4	23	4.233	.047*
Co	8.83	2.60	3	112			8.39	2.43	5	13		
	Searching											
	Task1						Task2					
	M	SD	min	max	F	p	M	SD	min	max	F	p
Ex	12.11	5.27	3	22	7.247	.011*	9.61	7.46	1	26	.607	.441
Co	8.17	4.66	2	19			8.1	3.36	4	14		

Note:*p<.05, Ex=experimental group, Co=control group

which can be interpreted that relevance judgment was hindered because of the burden of MRT.

2) Searchers' Perceived Emotions

For searcher's perceived emotions, participants often expressed their feelings about their level of search performance to see if they are searching the relevant information about the search task. If they felt that they performed the search poorly, in which there was no relevant information that they need, they tended to become frustrated and stressed. For example:

"I was frustrated when my first search was failed to find relevant information about the task. There is no information that I need so I become frustrated and stressed. Whew… there is no relevant information that I need. I don't know anything about what to do and how to search. Am I right? Is this right? Am I doing the right thing? I am not sure what I did was right."

These excerpts show that participants show their emotional instabilities when they are not sure of their search performance, either good or bad, which could lead to the state of frustration and stress. However, there was no significant difference in terms of participants' confidence and confusion between the two conditions. Confidence and confusion are more likely to be related to the level of search performance in that participants might be more confident or get confused

〈Tab. 4〉 Summary of perceived emotions

	Confidence											
	Task1						Task2					
	M	SD	min	max	F	p	M	SD	min	max	F	p
Ex	2.56	1.76	1	8	.837	.367	3.22	1.90	1	7	.312	.580
Co	2.11	1.08	1	4			2.89	1.68	1	7		
	Frustration											
	Task1						Task2					
	M	SD	min	max	F	p	M	SD	min	max	F	p
Ex	3.33	1.41	1	6	46.970	.000*	3.22	2.26	1	8	8.988	.005*
Co	0.611	0.92	0	3			1.28	1.56	0	5		
	Stress											
	Task1						Task2					
	M	SD	min	max	F	p	M	SD	min	max	F	p
Ex	4.83	2.53	1	9	11.264	.002*	3.67	1.81	1	7	5.414	.026*
Co	2.56	1.39	0	6			2.39	1.46	1	6		
	Confusion											
	Task1						Task2					
	M	SD	min	max	F	p	M	SD	min	max	F	p
Ex	4.50	2.81	1	11	.064	.802	4.89	2.72	1	10	.272	.606
Co	4.78	3.72	0	13			4.33	3.61	0	12		

Note:*p<.05, Ex=experimental group, Co=control group

after they perform well or poorly for the search task. Clore and Huntsinger (2007) stated that emotions influence on human judgment and thoughts, in which positive affect is found to promote but negative affect to inhibit many phenomena.

As can be seen in Table 4 below, the ANOVA results showed that there was a significant difference in terms of frustration and stress in searcher's perceived emotions. For task1, the experimental group searchers scored M=3.33 (SD=1.41) for frustration and M=4.83 (SD=2.53) for stress, while the control group searchers scored M=0.61 (SD=0.92) for frustration and M=2.56 (SD=1.39) for stress. For task2, the experimental group searchers scored M=3.22 (SD=2.26) for frustration and M=3.67 (SD=1.81) for stress, while the control group searchers scored M=1.28 (SD=1.56) for frustration and M=2.39 (SD=1.46) for stress.

3) Searchers' Perceived Efforts

From the excerpts it seemed that participants tended to often work harder than usual during information search task.

"It takes more time than usual. It looks like the quality of queries lowers. I am visiting this page to check relevance. I am focusing the materials more than usual. It's like I put more efforts than usual. I had to look at more pages than usual."

These excerpts suggest that requiring more mental load to perform a search task may encourage them to put more time to search, make a qualified query, visit more pages, or put more efforts to finish the search.

It seemed that being required to both searches for relevant information and do another task (i.e., MRT) at the same time can sometimes be taxing. However, when looking at the ANOVA results, there is a significant difference only in work hard subcategory of efforts among them. No significances were found among other subsections such as time spent, queries, and pages visit. However, there was a significant difference in work-hard subsection between the two groups with mental demand manipulation incurring more work to do for the experimental group (See Tab. 5).

<Tab. 5> Summary of perceived efforts

	Work-hard subsection											
	Task1						Task2					
	M	SD	min	max	F	P	M	SD	min	max	F	P
Ex	1.89	1.64	0	5	12.737	.001*	1.33	1.53	0	5	7.692	.009*
Co	0.39	0.70	0	2			0.22	0.73	0	3		

Note:*p<.05, Ex=experimental group, Co=control group

4) Searchers' Perceived Performance

In this category, we examined both on searchers' positive and negative verbal expressions associated with performance and satisfaction. For example:

"I wasn't 100 percent sure if I did find the relevant information (e.g., negative expression in performance). I think I found the relevant information so my performance was okay (e.g., positive expression in performance). I was levied at first due to the code memory but I performed the search task well. I think I can quit at this moment with this information"

As for the searcher's perceived performance, there was a significant difference between the two groups in regard to positive expressions in the performance subsection for both tasks and positive expressions in the satisfaction subsection for task1 as shown in Table 6. This could mean that mental demand has negative effects on searcher's perceived performance towards searching for relevant information.

In this category, searcher's perceived performance came along with relevance judgment. One explanation is that when a searcher found relevant information for the search task, they typically satisfied their performance that led to a positive expression of performance or satisfaction. Moreover, the results showed that only positive expressions in performance and satisfaction subsections were found to be significant (See Table 6), which could imply that searchers are more likely to express their performance at a time that they feel satisfied with search results at the end of the search task.

〈Tab. 6〉 Summary of perceived performance

	Performance (Positive)											
	Task1						Task2					
	M	SD	min	max	F	P	M	SD	min	max	F	P
Ex	3.33	2.06	1	7	6.186	.018*	3.28	2.67	0	9	4.262	.047*
Co	5.44	2.96	1	11			5.28	3.12	1	12		
	Satisfaction (Positive)											
	Task1						Task2					
	M	SD	min	max	F	P	M	SD	min	max	F	P
Ex	0.44	0.70	0	2	8.156	.007*	0.33	0.59	0	2	1.889	.178
Co	2.50	2.29	0	8			0.67	0.84	0	3		

Note:*p<.05, Ex=experimental group, Co=control group

2. NASA-TLX Cognitive Load

Because mental demand manipulation was made in this study, we had to check to see if it was successfully added to the participants in the experimental group. The pilot test confirmed that manipulated mental demand was checked properly between the two groups. An ANOVA was conducted to examine whether there were differences between the two groups in terms of each component of the NASA-TLX, namely, mental demand, physical demand, temporal demand, effort, and frustration. NASA-TLX has been used to measure cognitive load in many areas (Pass et al. 2003). As can be seen in Table 7 below, the results showed that significant difference was found in the component of mental demand (F=0.048, p=.048 in task1 and F=6.252, p=.017 in task2), performance (F=6.666, p=.014 in task1 and F=7.081, p=.012 in task2), and frustration (F=8.248, p=.007 in task1 and F=4.766, p=.036 in task2) both in task1 and task2.

No significant difference was found in physical demand between the two groups incurring no more workload. Although temporal demand to perform every task was not significant, a

<Tab. 7> Summary of NASA-TLX scores

	Mental demand							
	Task1				Task2			
	M	SD	F	P	M	SD	F	P
Ex	5.00	1.02	4.208	.048*	5.22	1.56	6.252	.017*
Co	4.17	1.38			4.00	1.37		
	Physical demand							
	Task1				Task2			
	M	SD	F	P	M	SD	F	P
Ex	3.50	1.20	.019	.890	3.83	1.47	.051	.822
Co	3.56	1.20			3.94	1.47		
	Temporal demand							
	Task1				Task2			
	M	SD	F	P	M	SD	F	P
Ex	5.33	1.46	2.092	.157	5.28	1.60	4.643	.038*
Co	4.61	1.54			4.00	1.94		
	Performance							
	Task1				Task2			
	M	SD	F	P	M	SD	F	P
Ex	4.00	0.84	6.666	.014*	3.72	1.32	7.081	.012*
Co	4.94	1.30			4.94	1.43		
	Effort							
	Task1				Task2			
	M	SD	F	P	M	SD	F	P
Ex	4.28	0.96	6.585	.015*	4.67	1.33	2.720	.108
Co	3.50	0.86			4.00	1.08		
	Frustration							
	Task1				Task2			
	M	SD	F	P	M	SD	F	P
Ex	5.50	1.29	8.248	.007	5.39	1.42	4.766	.036*
Co	4.33	1.14			4.22	1.77		

Note:*p<.05, Ex=experimental group, Co=control group

significant difference was found especially in task2 only, the most difficult task. It is interesting to see that the mean values to almost all dimensions were higher in the experimental group than the control group except performance dimension (i.e., 1=low to 7=high performance). According to the means, the experimental group reported the highest mean of cognitive load, whereas the control group demonstrated the lowest mean of cognitive load during information search.

3. Correlations between Each of the Mental State Indicators

1) Correlation in thoughts category

Pearson correlation coefficients were computed to examine whether there were relationships among subsections in each category. There were no significant relationships among subsections in thoughts category for both groups in task1. However, in task2, there were significant relationships between relevance judgment and thinking (r=.681, p=.002), searching and thinking (r=.485, p=.041), searching and relevance judgment (r=.717, p=.001) for the experimental group, and between searching and remembering (r=.504, p=.033) for the control group. A possible reason for this finding is that task2 is considered more difficult to perform than task1 so that the participants in the experimental group were not able to reduce cognitive load and thus obtain more work to finish the search task.

〈Tab. 8〉 Correlation coefficients among subsections in thoughts category

Experimental group in task1					Control group in task1				
	1	2	3	4		1	2	3	4
1.tk	1				1.tk	1			
2.rm	.304	1			2.rm	.277	1		
3.rj	-.051	.152	1		3.rj	.072	.276	1	
4.sc	.060	-.010	.266	1	4.sc	.456	.380	.041	1
Experimental group in task2					Control group in task2				
	1	2	3	4		1	2	3	4
1.tk	1				1.tk	1			
2.rm	-.330	1			2.rm	-.032	1		
3.rj	.681**	-.198	1		3.rj	.108	.409	1	
4.sc	.485*	.179	.717**	1	4.sc	.145	.504*	.452	1

Note:*p<.05, **p<.01, tk=thinking, rm=remembering, rj=relevance judgement, sc=searching

2) Correlation in Emotions Category

The results showed that there were positive relationships between the following pairs of stress and frustration in task1, and stress and frustration, and confusion and confidence in task2 for the experimental group, while there were significant relationships between the pairs of confusion and confidence in task1, and confusion and confidence, and confusion and frustration, and confusion and stress in task2 for the control group. It appears that stress and frustration closely relate to each other for the experimental group, whereas confidence and confusion closely relate each other for the control group.

〈Tab. 9〉 Correlation coefficients among subsections in emotions category

Experimental group in task1					Control group in task1				
	1	2	3	4		1	2	3	4
1.cfd	1				1.cfd	1			
2.frs	.205	1			2.frs	-.099	1		
3.str	-.044	.576*	1		3.str	-.108	.128	1	
4.cfs	.452	.399	.368	1	4.cfs	.505*	.392	-.078	1
Experimental group in task2					Control group in task2				
	1	2	3	4		1	2	3	4
1.cfd	1				1.cfd	1			
2.frs	.209	1			2.frs	.362	1		
3.str	-.154	.593**	1		3.str	.179	.363	1	
4.cfs	.575*	.320	.165	1	4.cfs	.628*	.515*	.587*	1

Note:*p<.05, **p<.01, cfd=confidence, frs=frustration, str=stress, cfs=confusion

3) Correlation in Efforts Category

The results showed that there were negative relationships between the pairs of query made and time spent in task1, and pages visit and query made, and work hard and pages visit in task2 for the experimental group. On the other hand, there was a positive relationship between the pairs of pages visit and query made in task2 for the control group.

〈Tab. 10〉 Correlation coefficient among subsections in efforts category

Experimental group in task1					Control group in task1				
	1	2	3	4		1	2	3	4
1.ts	1				1.ts	1			
2.qm	-.517*	1			2.qm	.425	1		
3.pv	-.194	.460	1		3.pv	.175	.406	1	
4.wh	.-.052	-.539	.059	1	4.wh	-.282	-.425	-.246	1
Experimental group in task2					Control group in task2				
	1	2	3	4		1	2	3	4
1.ts	1				1.ts	1			
2.qm	.241	1			2.qm	.261	1		
3.pv	-.035	-.608*	1		3.pv	.110	.718**	1	
4.wh	-.184	.207	-.497*	1	4.wh	.000	.016	.070	1

Note:*p<.05, **p<.01, ts=time spent, qm=query made, pv=pages visit, wh=work hard

4) Correlation in Performance category

The results showed that there was a positive relationship between the pairs of performance and satisfaction in task2 only for the control group (r=.867, p=.000), which implies that user satisfaction may depend on performance without any extra mental demand. It is argued that the relationship between performance and overall satisfaction judgment is much complex. In this study, this might be due to the fact that the participants might not fully present their verbal expressions about the level of their search performance as well as search satisfaction during concurrent think aloud protocol in a course of a search task. Using a quantitative approach such as a survey questionnaire could reveal more direct results about the relationship between the variable pairs of performance and satisfaction.

Ⅵ. Conclusion and Implications

The aim of this paper was to present mental states of searchers when they have mental demand as well as assessment of cognitive load on the given search tasks. We have done so by conducting an exploratory user study that employed a dual-task based on MRT effect as a technique for manipulating mental demand on web search tasks. Accordingly, the first major practical contribution of the present research is that it provides a big picture of information task search process about what a searcher experiences during task performance with what in mind. Understanding dynamic changes of mental state is important and useful for increasing our knowledge about user search behavior of information search process.

We found four categories of mental state were found to vary by a search task. For participants' perceived thoughts, the results indicated that the participants with mental demand did express thoughts more often than those without mental demand in frequency of thoughts term occurrences based on the subsections of coding scheme. For participants' perceived emotion, there was a significant difference in terms of frustration and stress in searcher's perceived emotions. For participants' perceived efforts, there was a significant difference in work-hard subsection between the two groups with mental demand manipulation incurring more work to do for the experimental group. As for the searcher's perceived performance, there was a significant difference between the two groups in regard to positive expressions in the performance subsection for both tasks and positive expressions in the satisfaction subsection for task1. The emerging picture presented here provides, in our opinion, a very important stepping stone towards a better design of usability and systems.

Another contribution to user interface and information search systems is that the searcher experience with mental demand was found to vary consisting of four different dimensions, namely, thoughts, emotions, efforts, and performance. The complexity of information search systems along with the complex search tasks may contribute to extra mental demand and result in giving up a task. This research opens areas for further developments in user interface for information search systems to examine how to design a positive search experience.

Ⅶ. Limitations and Future Direction

Memory and recall task used in this study was somewhat considered weak. It was reasonable to induce extra mental demand memory load but for future study a stronger mental demand manipulation may yield more interesting mental states. In this study, we used indirect and subjective methods to assess mental states and cognitive load. For future research, direct and objective methods are expected to draw more interesting effects.

A concurrent think-aloud protocol may make a searcher feel uneasy during information search task and it could have increased extrinsic cognitive load and may already have required germane cognitive load in addition to a secondary task (i.e., MRT). Therefore, there might have more room for misinterpretation of information, given that much information should be being processed.

In the present study, we attempted to describe mental states and cognitive load in information search process. It is important to acknowledge that this study is a starting point for mental states from the perspective of psychological description in the area of information retrieval. In the future research, it is also important to continue this study and extend the scope to elsewhere such as collaborative information search for an example.

Another limitation concerns the small sample size (36 participants), which may impact the generalization of the results and external validity of this exploratory study to the whole population of the university students.

Further studies, using physiological methods, in addition to think-aloud protocol, are needed to further examine a searcher's body reactions to mental demand. The study has also implications for future information retrieval research. The results of this study in an experimental setting warrant a replication in an information retrieval study, in particular as mental demand and associated variables such as mental states including thoughts, emotions, effort, and performance.

References

Annett, J. 2002. "Subjective rating scales: science or art?" *Ergonomics,* 45(14): 966-987.

Bandura, A. 1982. "Self-efficacy mechanism in human agency." *The American Psychologist,* 37(2): 122-147.

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

Blumberg, S. J. 2000. "The White Bear Suppression Inventory: Revisiting its factor structure." *Personality and Individual Differences,* 29: 943−950.

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

Bystrom, K. 2002. "Information and Information Sources in Tasks of Varying Complexity." *Journal of the American Society for Information Science and Technology,* 53(7): 581-591.

Byström, K. 2005. "Conceptual framework for tasks in information studies." *Journal of the American Society for Information Science and Technology,* 56(10): 1050-1061.

Chen, I. J. and C. C. Chang. 2009. "Cognitive Load Theory: An Empirical Study of Anxiety and Task Performance in Language Learning." *Electronic Journal of Research in Educational Psychology,* 7(2): 729-746.

Clore, G. and J. R. Huntsinger. 2007. "How emotions inform judgment and regulate thought." *Trends in Cognitive Science,* 11(9): 393-399.

Craik, F. and Lockhart, P. 1972. "Levels of processing: A framework for memory research." *Journal of Verbal Learning and Verbal Behavior,* 11: 671-684.

Garrison, D. R., Cleveland, M., Kool, M. and J. Kappelman. 2006. "Revisiting methodological issues in the analysis of transcripts: negotiated coding and reliability." *Internet High Education,* 9(1): 1−8.

Gimino, A. 2002. "Factors that influence students' investment of mental effort in academic tasks: a validation and exploratory study." Doctoral Dissertation.

Gwizdka, J. 2008. “Cognitive Load on Web Search Tasks.” *Proceedings of Workshop on Cognition and the Web,* 83-86

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

Gwizdka, J. 2013. “Searchers switch tactics under increased mental load.” P*roceedings of the Association for Information Science and Technology,* 50(1): 1-3.

Howells, F. M., Stein, D. J. and V. A. Russell. 2010. “Perceived mental effort correlates with changes in tonic arousal during attentional tasks.” *Behavioral and Brain Functions,* 6(39): 1-15.

Kasl, S. V. 1996. “The influence of the work environment on cardiovascular health: A historical, conceptual, and methodological perspective.” *Journal of Occupational and Health Psychology,* 1: 42–56.

Kirschner, F., Paas, F. and P. A. Kirschner. 2009. “A cognitive load approach to collaborative learning: United brains for complex tasks.” *Educational Psychology Review,* 21: 31-42.

Kirschner, F., Paas, F. and P. A. Kirschner. 2011. “Superiority of collaborative learning with complex tasks: A research note on an alternative affective explanation.” *Computers in Human Behavior,* 27: 53-57.

Kirschner, P. A. 2002. “Cognitive load theory: Implications of cognitive load theory on the design of learning.” *Learning and Instruction,* 12(1): 1-10.

Kuhlthau, C. C. 1991. “Inside the search process: Information seeking from the user’s perspective.” *Journal of the American Society of Information Science,* 42(5): 361-371.

NASA, NASA TLX: Task Load Index. <http://humansystems.arc.nasa.gov/groups/tlx> [cited 2017. 8. 28].

Nielsen Norman Group. Usability 101: Introduction to Usability. <http://www.nngroup.com/articles/usability-101-introduction-to-usability> [cited 2017. 8. 28]

Niculescu,, A., Cao, Y., Vandijk, B. and A. Nijholt. 2010. “Measuring Stress and Cognitive Load Effects on the Perceived quality of a multimodal dialogue system.” *Proceedings of Measuring Behavior:* 453-455.

Oosterwijk S., Lindquist, K. A., Anderson, E., Dautoff, R., Moriguchi, Y. and L. F. Barrett. 2012. "States of mind: Emotions, body feelings, and thoughts share distributed neural networks.” *Neuroimage,* 62(3): 2110-2128.

Paas, F. and J. Merriënboer. 1993. “The efficiency of instructional conditions: An approach to combine mental effort and performance measures.” *Human Factors,* 35: 737-743.

Paas F., Van Gerven P. W. M., and H. K. Tabbers. 2005. *The cognitive aging principle in multimedia learning.* In Mayer R. E. (Ed.), The Cambridge handbook of multimedia learning. 339-354. New York: Cambridge University Press.

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

Peterson, R. D., Klein, J., Donnelly, R. and K. Renk. 2009. "Predicting psychological symptoms: The role of perceived thought control ability." *Cognitive Behavior Therapy,* 38(1): 16-28.

Pew Research Cente. Three Technology Revolutions. <http://www.pewinternet.org/Trend-Data/Online-Activites-Total.aspx> [cited 2017. 8. 28].

Pu, P. H. and K. Pratyush. 2004. "Evaluating example-based search tools." *Proceedings of the 5th ACM conference on Electronic commerce,* 208-217.

Salomon, G. 1984. "Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions." *Journal of Educational Psychology,* 76: 658.

Seibert, P. S. and H. C. Ellis. 1991, "Irrelevant thoughts, emotional mood states, and cognitive task performance." *Memory Cognition,* 19(5): 507–513.

Sweller, J. 1988. "Cognitive load during problem solving: Effects on learning." *Cognitive Science,* 12(2): 257-285.

Sweller. J. 2010. "Element interactivity and intrinsic, extraneous, and germane cognitive load." *Educational Psychology Review,* 22(2): 123-138.

Trzepacz, P. T. and R. W. Baker. 1993. "Mental status examination." Oxford, U.K.: Oxford University Press. p.202.

Vakkari, P. 1999. "Task complexity, problem structure and information actions: integrating studies on information seeking and retrieval." *Information processing & management,* 35(6): 819-837.

Vakkari, P. and P. Pertti. 2001. "Changes in search tactics and relevance judgements when preparing a research proposal: a summary of the findings of a longitudinal study." *Information retrieval,* 43(4): 295-310.

Wilson, T. D. 2006. "On user studies and information needs." *Journal of Documentation,* 62(6): 658-670.

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

Zhang, Y and Gwizdka, J. 2014. "Effects of tasks at similar and different complexity levels." *Proceedings of the Association for Information Science and Technology,* 51(1): 1-4.