Dual-Task Performance as a Measure of Mental Effort in Searching a Library System and the Web

Yong-Mi Kim, Corresponding Author

School of Information, University of Michigan, 1075 Beal Ave., Ann Arbor, MI 48109-2112.kimym@umich.edu

Soo Young Rieh

School of Information, University of Michigan, 1085 South University Ave., Ann Arbor, MI 48109-1107.rieh@umich.edu

This paper examines a dual-task method for the assessment of mental effort during online searching, having the users engage in two tasks simultaneously. Searching was assigned as a primary task and a visual observation was set up as a secondary task. The study participants were asked to perform two searches, one on the Web and the other in a web-based library system. Perceived search difficulty and mental effort for searching on the two types of systems were compared through participant self-reports, dual-task performance, and search log analysis. After the searches were completed, the subjects reported that library searching was more difficult to conduct and they had to concentrate more than when Web searching. However, the results of dual-task performance do not reveal much difference in mental effort or concentration during searches in the two systems. Rather, they invested mental effort differently when viewing search results and reading retrieved documents. The findings indicate that a dual-task method provides a useful technique to measure mental effort in online searching, and it has a great potential to be used to measure other aspects of information retrieval such as task complexity and multitasking information behavior.

Introduction

While there have been numerous studies of users of information retrieval systems in the experimental lab environment, there has not been similar variety in the methods employed to measure user perceptions and behaviors in searching. Methods employed to date have been restricted to observation and user self-reports by means of questionnaires, think-aloud protocols during searches, and post-search interviews. There has been little effort in developing or exploring alternative methods of data collection during the search process. This paper examines the application of a dual-task method as a way of measuring mental effort or concentration in online searching.

This research compares online searching in a web-based academic library system and a web search engine. These two systems were chosen because they are the most popular information retrieval systems. Online public access catalogs (OPACs) were once the most widely used information retrieval system (Borgman, 1996), but web search engines have become the systems of choice for most users (Pew Internet Project, 2005). Despite the fact that web-based library systems and web search engines provide different kinds of resources, services, and interfaces, people's searching behavior across these two systems seems to be fairly consistent in its simplicity. Over the years, Spink and her colleagues have analyzed the transaction logs of web search engines and discovered how people conduct web searches: sessions are brief (2.4 queries per user) with short queries (2.5 terms per query), viewing 1.7 pages per query. They rarely use Boolean operators and advanced features for web searching (Silverstein, Henzinger, Marais, & Morica, 1999; Spink, Jansen, Wolfram, & Saracevic, 2002; Spink & Jansen, 2004). These characteristics – short queries, brief sessions, and infrequent use of Boolean operators and advanced system features – are also typical of the online searches that people conduct in other information retrieval systems such as online catalogs and CD-ROM databases (Cooper, 2001; Drabenstott & Vizine-Goetz, 1994; Sutcliffe, Ennis, & Watkinson, 2000).

This leads to an interesting research problem: why do people put so little effort into online searching? The related question would be: do the amounts of effort that people put into searching on the web and searching in library systems differ? In a previous study, Rieh (2005) examined these research questions employing the methods of pre-search questionnaires, post-search questionnaire, and post-system interviews along with collection of search logs. The findings of that study indicate that study subjects perceived that web searching would be much easier than library system searching, these differential perceptions then influencing the level of thought, concentration, and effort they put into online searching. Limitations of the previous study include that although the search logs were analyzed the data analysis relied primarily on the subjects' self-reports focusing on their perceptions of search difficulty and mental effort.

This study addresses a similar set of research questions by employing a different method to measure mental effort in online searching. In addition to asking the subjects to directly assess their levels of effort and searching difficulty, the researchers employed a dual-task method by having subjects engage in two tasks simultaneously. Dual-task methods have been widely used in experimental psychology to measure cognitive load; however, it has not been extensively used in IR experiments (Dennis, Bruza, & McArthur, 2002). This research was designed such that searching on the web or in the library system was the primary task and visual observation was the secondary task.

Using the dual-task method, the present study brings two strands of research together: comparison of searching behavior in information retrieval systems, and measurement of mental effort invested into online searching. The specific research questions addressed in this study include:

- Does people's investment of effort into online searching differ on the web and the web-based library system?
- Do people invest differential effort depending on the types of search activities that they engage in?
- Does a dual-task method provide a reliable way of measuring mental effort or other attributes of searching performance?

Mental Effort

According to Salomon (1984), mental elaborations can range from the automatic and effortless to the more effort-demanding, controlled ones. Salomon (1981) introduced the notion of amount of invested mental effort (AIME), defining it as the number of non-automatic mental elaborations necessary to solve a problem. He explains that learners will invest greater effort in processing material when they encounter complex, ambiguous, incongruent, or novel stimuli that cannot easily be accounted for by their existing mental schema. In other words, AIME can be expected to decrease when learners perceive the encountered material to be easy, warranting little investment of mental effort.

Researchers have noted that the terms mental effort, attention, concentration, use of cognitive capacity, and mental workload all refer to similar concepts (Cennamo, 1993). These are all related to an increase in the cognitive resources devoted to processing stimuli (Britton, Muth, & Glynn, 1986). Cennamo argues that not only characteristics of the media but also characteristics of the task and the learner influence preconceptions of the media, which then impacts the amount of mental effort invested.

A number of studies examine how people's perceptions of television viewing influence the level of effort they put into that activity (Cennamo, 1993; Krendl, 1986; Salomon, 1984). Comparing learners' perceptions of television and print media, Salomon found that television is perceived as being easier than print, and that less mental effort is invested in learning from the former than the latter. As a result, less is learned from television than from print. Salomon concluded that learners' preconceptions of television as an easy medium influence the mental effort they expend in processing a television lesson, and that the amount of mental effort learners invest in a lesson influences learning achievement (Salomon, 1984).

Parallels can be drawn between television viewing and web searching. As the web is used for various purposes such as entertainment, social interaction, and obtaining news, people may perceive the web as being easy (Pew Internet Project, 2004). In contrast, library systems may be perceived as "serious" media because often libraries are used when doing schoolwork or work-related tasks. Therefore, people might feel that searching on the library system requires more effort and concentration than searching on search engines (Fast & Campbell, 2004).

Dual-Task Method

Researchers from many different fields have been interested in assessing mental effort for a variety of tasks. Cennamo (1993) summarizes methods for assessing mental effort into three categories: opinion measures, dual-task techniques, and physiological measures. Opinion measures include a variety of self-report measures used to assess mental effort. Opinion measures assume that "the investment of effort is a voluntary process that is under the control of the individual and as such is available for introspection" (Cennamo, p. 36). Dual-task techniques encompass a range of methods whereby the study subject is assigned two tasks to be performed at the same time: a primary task such as reading a passage, viewing a television program, or doing a web search, and a secondary task such as responding to a tone or light flash, or performing mental arithmetic. The assumption is that there is a limit to people's cognitive capacity, so that "when a great deal of cognitive capacity is consumed by the primary task, less capacity is available to devote to the secondary task" (Cennamo, p. 37). Lastly, physiological measures, such as heart rate, pupil dilation, or, more recently, brain activity measurements have been used to assess mental effort. The assumption is that physiological changes occur with changes in the demand for mental effort (Iani, Gopher, & Lavie, 2004; O'Donnell & Eggemeier, 1986).

The dual-task method is widely used in experimental psychology and human factors research (Pashler, 2000; Verwey & Veltman, 1996), though not as prevalently in IR research (Dennis, Bruza, & McArthur, 2002). Two variations of dual-task methodology are commonly used in experimental research (Brünken, Steinbacher, Plass, & Leutner, 2002). In one, experimenters study the reduction in performance of the primary task in the dual-task condition, while in the other they study the reduction in performance of the secondary task in the dual-task condition, in each case as compared to the single-task condition. Studies assessing cognitive load look at the reduction of secondary task performance, which is most often measured using reaction time (interval of time between stimulus and user reaction) and error rate (failure to respond to a stimulus or to respond correctly).

In designing experiments using dual-task methodology to assess mental effort, a number of factors affecting dual-task performance must be considered. Practice or familiarity with a task is one factor. With practice or repetition of a task the mental effort required may decrease, with repeated trials possibly having the effect of training subjects for the secondary task. This results in secondary task performance that has little to do with the difficulty of the primary task. Thus the experimenter must be careful about determining the frequency with which stimuli are presented in a session as well as about the length of the session itself. The timing of stimuli (or the interval of time between stimuli) must also be considered because if the stimuli occur too regularly subjects may then start to anticipate them, consequently affecting reaction time and accuracy (Grosjean, Rosenbaum & Elsinger, 2001). Another factor is the modality (e.g., visual or auditory) employed in the primary and secondary tasks. For example, primary and secondary tasks will interfere more with each other if they both require responding to visual stimuli, as compared to a scenario in which one task requires responding to visual stimuli and the other to auditory stimuli (Eysenck, 2001). If the primary and secondary tasks are too dissimilar there may be no significant difference in task performance in single-task and dual-task conditions.

Researchers interested in using dual-task methods must give careful consideration to the experimental apparatus required. Reaction times must be measured in milliseconds, which is not feasible with commonly used commercial products for recording subject responses. This lack may require researchers to develop or build their own measurement apparatus. The particular combination of primary and secondary tasks a researcher selects for use may require additional development or modification of experimental equipment.

Research Design

Participants

Nine undergraduate students attending the University of Michigan participated in the study. They were recruited through a web site that solicits volunteers for paid experiments at the University of Michigan, as well as through requests for participation delivered via graduate student instructors in undergraduate courses in the School of Information, the Department of Psychology, and the Department of Political Science. The recruitment text solicited the participation of students who had been assigned a term paper, course assignment, or project requiring the use of library resources. The demographic profile of the study subjects is summarized in Table 1.

Table 1. Summary of Participant Profiles

	Class Level	Major	Gender	Topic
1	Senior	CS	M	Mars Rover
2	Freshman	Undecided	M	Biometrics and e-voting
3	Freshman	Economics	M	British empire in India
4	Freshman	Undecided	M	E-voting
5	Sophomore	Math	F	Buddhism in China
6	Senior	Psychology	F	Political psychology of negative campaigning
7	Senior	Music	M	History of Javanese Gamelan music
8	Freshman	Biology	F	Effects of altruism on individuals
9	Freshman	Biochemistry	M	Life of St. Benedict

Experimental Design

A within-subject design was selected for two reasons: (1) the large individual differences expected in reaction times; (2) the obtaining of the subject's own assessment of the comparative difficulty of searching in two different systems.

Primary task

Each study subject performed two searches, one in each system. For the library condition, subjects searched from the University of Michigan Library's main page (http://www.lib.umich.edu). For the web search condition, subjects were allowed to start from the web site they normally use for such tasks.

Secondary task

A visual observation task was selected as the secondary task. As mentioned previously, one concern in dual-task experiments is that if the primary and secondary tasks are too dissimilar there may not be a significant difference in task performance under single-task and dual-task conditions. Some previous experimental results employing dual-task methods indicated that tasks of different modalities might prove to be too dissimilar (Brünken, Steinbacher, Plass, & Leutner, 2002; Dennis, Bruza, & McArthur, 2002) so a visual task was selected for this IR experiment. A relatively simple task was selected to focus task performance on the primary task. A complex secondary task runs the risk of the experimental subject focusing on that task and not the primary one.

A small window was displayed on the upper right area of the screen, to the right of the browser window (see Figure 1). After a random period of 45-75 seconds, the window changed color, either from grey to red or red to grey. Subjects were asked to press the escape key of the keyboard as soon as they perceived the color change. Software recorded the time lapse between the color change and the pressing of the escape key. The secondary task application was implemented in Java 1.4.2, and subject reaction times were recorded by configuring a freeware program called Global Hotkey, which allows certain actions to be mapped to keys on the keyboard.

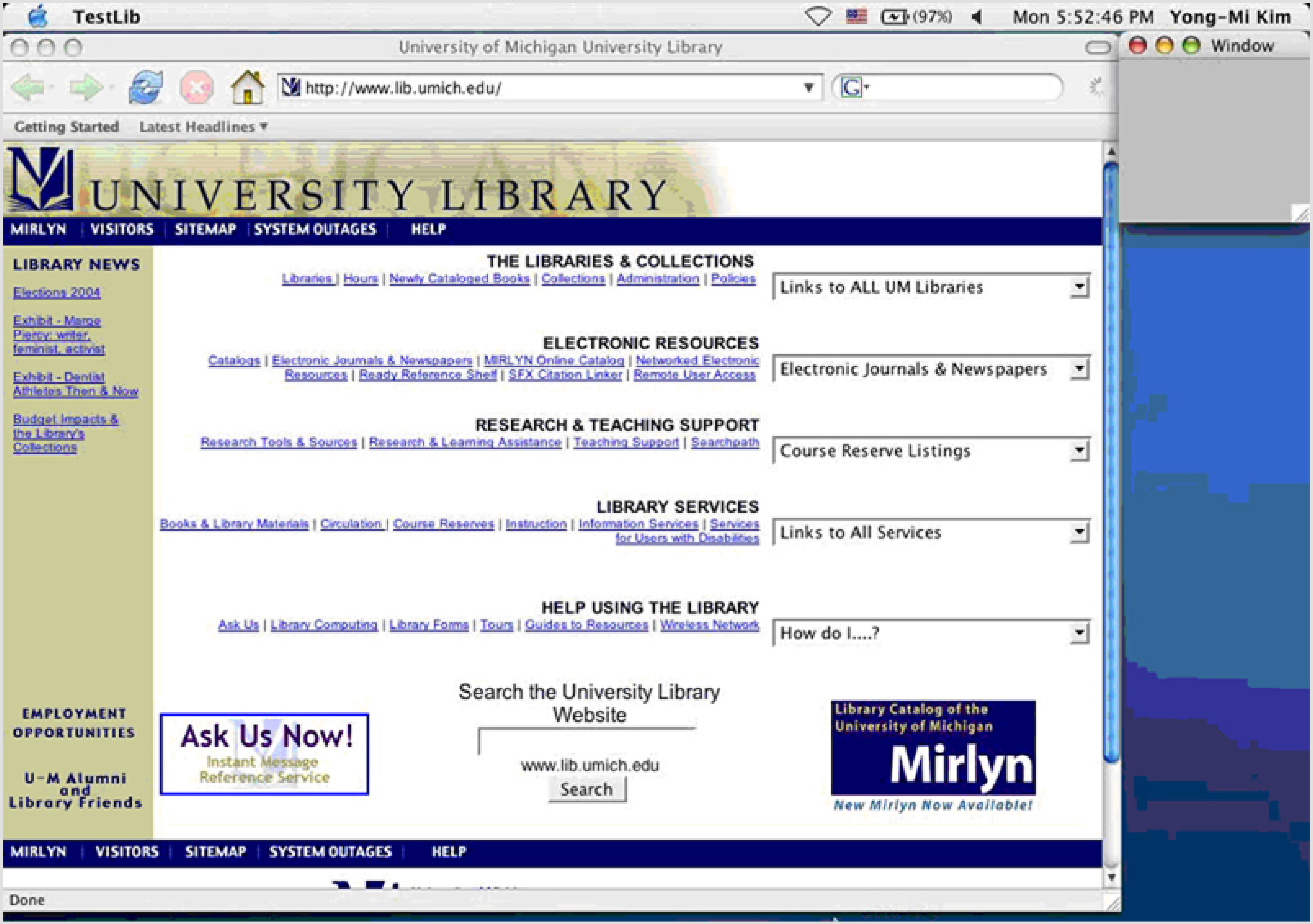

Figure 1: Dual-task condition – library search

Experimental Apparatus

Study subjects carried out their searches using the Firefox 1.0 browser running on an Apple G4 iBook with a 12” screen and OS X 10.3.6. Searches were captured as QuickTime movies using Snapz Pro X 2, a Macintosh-based screen capture program. Post-search interviews were recorded on audiotape. The experiments took place in November and December of 2004.

Procedures and Tasks

Study subjects signed a consent form and completed a background questionnaire prior to starting their searches. The background questionnaire asked for demographic information, search experience, perceptions of the level of difficulty in searching the web and library systems, and confidence in their ability to search in both systems. Subjects were given a brief demonstration on use of the Firefox browser. Printed instructions were provided, which were also read aloud before each search task. They were told they would be limited to 15 minutes on each system. Finally, they were instructed to perform their searches as they normally would when they perform searches for their term paper topics.

Each study subject was presented with three conditions (single-task, dual-task/library search, and dual-task/web search) in one of four different orders (see Table 2). In the single-task condition the subject was shown a blank browser window along with the secondary task window. The single-task condition was included in our experimental design to obtain a baseline measure of each subject’s reaction time. Both the library and web searches were dual-task conditions in which the subject had to respond to the secondary task while carrying out their searches.

Table 2. Treatment conditions

Treatment	Condition 1	Condition 2	Condition 3
1	Secondary task only	Dual-task library	Dual-task web
2	Secondary task only	Dual-task web	Dual-task library
3	Dual-task library	Dual-task web	Secondary task only
4	Dual-task web	Dual-task library	Secondary task only

The on-screen activity was captured using Snapz Pro X 2. The researcher observed the study subjects as they searched and took notes which were later referred to during data analysis. After completion of each search, subjects completed a post-search questionnaire that included questions on the difficulty they experienced during their search, as well as the degree of concentration required in their searching. The researcher also asked additional questions regarding the search, such as why a subject selected a particular database or whether the subject had received instruction in the advanced search functions they employed. At the end of both searches an exit interview was conducted in which subjects were asked to rate on a 0-10 scale a number of aspects regarding searching on the web and in the library system (e.g., the difficulties of using each system and how likely they would be to use each system for similar searches in the future). Subjects were also asked open-ended questions on what factors they thought contributed to successful and unsuccessful searches in each system. Their answers were recorded, transcribed, and analyzed.

Results

Self-reports

As presented in Table 3, it was found that before the subjects conducted their searches, they perceived that searching in the library system would be significantly more difficult than searching on the web ($t(7) = 1.95$, $p < 0.05$). They were also significantly more confident with web searching than library system searching ($t(7) = 1.92$, $p < 0.05$).[1] Table 4 shows the extent of effort, thought, and concentration subjects put into their searches under the web and library conditions as summarized from the post-search questionnaires. Significant differences were found for the perceived difficulty in starting the search and difficulty in understanding search results. The library system was perceived as demanding more concentration.

Table 3. Pre-task perceived difficulty and confidence

Questions	Web		Library	
	M	SD	M	SD
Estimated difficulty of search*	1.44	1.01	4.38	2.45
Confidence in ability to find information*	6.67	3.08	4.38	2.26

* $p < 0.05$; scale of 0 (not difficult, not confident) to 10 (very difficult, very confident)

Table 4. Comparison of difficulty and effort

Questions	Web		Library	
	M	SD	M	SD
Difficult to start the search*	1.44	1.94	5.0	2.74
Difficult to understand search results*	2.67	2.0	4.33	2.6
Difficult to make decisions of usefulness	5.0	2.24	5.78	2.39
Effort invested into searching	4.56	1.59	6.22	1.3
Thought put into searching	5.67	1.22	6.33	1.41
Concentration during search*	6.44	1.24	8.0	1.12

* $p < 0.05$; scale of 0 (not difficult, no effort, no thinking, no concentration) to 10 (very difficult, a great deal of effort, a great deal of thinking, a great deal of concentration)

Dual-task results

Depending on their task time for the dual-task conditions, subjects were presented with 9-15 color changes. Two measures of dual-task performance were initially defined: reaction time (RT), or the amount of time it takes a subject to respond to the visual stimulus; and miss frequency, or the number of times a subject fails to respond to the visual stimulus. In the experiments it was observed that some subjects hit the escape key even though there had been no visual stimulus, and these events were referred to as false alarms.

Table 5 displays the individual mean reaction times of the three conditions for all subjects[2]. The single-task condition had the lowest reaction times measured in milliseconds (M=602.78, SD=114.32), followed by the web search dual-task condition (M=3310.11, SD=2061.92) and the library search dual-task-condition (M=4837.78, SD=6515.63). Differences in reaction times between the single-task condition and the web search dual-task condition (t_8=3.945, p<0.005), and the single-task condition and the library search dual-task condition (t_8=1.95, p<0.05) were found to be statistically significant. The difference in reaction times in the library search dual-task condition and the web search dual-task condition was not statistically significant.

Table 5. Individual reaction times (in milliseconds)

Subject	Single-Task		Dual-Task Web		Dual-Task Library	
	M	SD	M	SD	M	SD
1	390	28	1521	127	1795	345
2	681	320	3579	6473	3901	4818
3	519	133	7974	10615	21942	22608
4	455	40	3956	5415	2639	3000
5	718	142	4441	2118	2156	433
6	709	267	1216	106	1484	341
7	651	286	2165	460	5143	4622
8	568	209	2367	2068	2431	1895
9	469	43	2020	410	2049	229
Average	602.78	114.32	3310.11	2061.92	4837.78	6515.63

Individual miss frequencies and false alarms are presented in Table 6. While the number of misses for the library search dual-task condition (M=2.78, SD=2.54) was higher than for the web search dual-task condition (M=1.78, SD=2.3), this difference was not statistically significant.

Table 6. Individual miss frequencies and false alarms

Subject	Dual-Task Web		Dual-Task Library	
	Miss	False Alarm	Miss	False Alarm
1	4	0	4	0
2	1	1	3	0
3	1	0	6	2
4	0	1	1	1
5	7	0	3	0
6	0	1	0	0
7	2	0	1	0
8	1	0	0	1
9	0	0	7	0
Average	1.78	0.33	2.78	0.44

Screen captures were examined to determine the state in which the secondary task errors (misses and false alarms) occurred (Table 7). Particularly striking is that for the web condition, no secondary task errors occurred when the subject was scanning search results, while the majority of errors occurred when the subject was reading a document. When the subject was scrolling on a page, errors occurred more frequently in the library condition than the web condition.

Table 7. Search state at time of error (miss frequencies and false alarms combined)

Search state	Web	Library
Reading	7	5
Scrolling	2	7
Entering query	2	2
Waiting for page	0	5
Search results	0	2
Other	8	8
Total	19	29

Search characteristics

Analysis of search logs showed that subjects spent more time reading documents than viewing search results in both their web and library searches, as presented in Table 8. But in comparing web and library searches, subjects spent more time viewing search results in the library system than on the web, while the opposite was the case for time spent reading documents.

Table 8. Time taken for activities (in minutes)

Activities	Web		Library	
	M	SD	M	SD
Time to complete task	14.35	2.12	14.43	1.87
Time to view search results	2.00	1.27	3.4	1.75
Time to read documents	9.24	3.34	5.45	3.09

Table 9. Search interactions

Search interactions	Web		Library	
	M	SD	M	SD
Clicks in total	28.78	17.42	47.89	18.60
Selection of search results	8.67	6.28	8.67	3.64
Documents read	13.56	8.62	12.44	5.59
Printing/email/saving	1.78	2.05	3.22	5.04

Table 9 shows that when performing their searches subjects had to click on more links in the library system than on the web.

Subjects also engaged in more instances of saving their search results in one form or another when using the library system. The library system offers several methods of retaining search results, such as emailing the document or saving the search result to a "basket" or "folder" active during a search session. While subjects used these mechanisms in the library system, they primarily printed the results or saved bookmarks when searching the web.

On average subjects used more terms, with a higher variance in the number of terms, for web searching than for library system searching (Table 10). There were more reformulations and advanced feature use in library queries. Several library system queries returned no results, while web queries always produced at least one result. Subjects did not rely solely on web search engines, but also used a site's search function if one was available.

Table 10. Query formulation behavior

Query formulation behavior	Web		Library	
	M	SD	M	SD
Terms in the first query	2.56	1.24	2.44	0.73
Terms in average	3.14	1.1	2.38	0.5
Reformulation	4.78	5.74	6.89	2.62
Advanced feature use	1.67	3.9	2.44	2.65

Discussion

1. Does people's investment of effort into online searching differ on the web and the web-based library system?

The results for Research Question 1 are mixed. Prior to conducting searches, subjects expected library system searches to be more difficult than web searches. In the exit interview, subjects reported that library system searches were indeed more difficult, especially in getting started with the search and understanding the search results. They also responded that they concentrated to a greater extent when they conducted searches in the library system than on the web.

However, there was no significant difference in secondary task performance for library and web searches. One explanation can be found in the disparity between performance and perception. A number of studies show that people are generally inaccurate in assessing their performance, both a priori and posteriori (Glenberg & Epstein, 1987; Maki & Berry, 1984). It can be argued that a similar disparity occurred in this experiment.

Another reason for such findings seems related to the type of task used for the study. In this experiment study subjects could search on a topic of their choice, with the restriction that it be related to an academic paper or research project. Notably, subjects who had an assignment soon after the experiment session were highly motivated in their searches. However, our previous study (Rieh, 2005) revealed when searching the web subjects perceived the research task (i.e., finding information for a term paper topic) to be much more demanding than the product task (i.e., finding reviews about a digital camera). As Cennamo (1993) argues, amount of invested mental effort (AIME) is not solely influenced by the characteristics of media. Characteristics of users and characteristics of task also influence the level of AIME. As subjects perceived the research task to be more demanding, they might have concentrated and invested effort at similar levels across the two different systems.

2. Do people invest differential effort depending on the types of search activities that they engage in?

The findings related to Research Question 2 are interesting. It was found that more secondary task misses were observed when subjects were reading documents than when they were viewing search results. This was particularly striking in the case of web searches, in which zero misses were observed when viewing search results. Observation of search behavior and post-search interviews indicate that subjects may not have been reading the search results but were scanning to see whether the keywords they had in mind were present, as well as to assess their proximity to each other.

While the total amount of time spent searching in library system and web searches was similar, allocation of time to different search activities took different patterns in each system. For example, subjects spent less time reading documents during library searches as compared to web searches. On the other hand, a larger chunk of library system search time was spent on navigation activities. The number of clicks in library system searching was much larger than in web searching, with subjects frequently clicking back several pages to return to the initial search query page. More query reformulation and use of advanced features was present in library system searching than in web searching. Advanced features such as Boolean operators and the use of wildcards (e.g., injur* instead of 'injury') were observed only in searches of online databases accessed through the library. In web searching, only one subject used advanced search features, consistently starting searches from Google's Advanced Search page.

3. Does dual-task methodology provide a reliable way of measuring mental effort or other attributes of searching performance?

There are a number of findings related to dual-task performance as an appropriate measure for mental effort in online searching. In terms of feasibility, this study did not find any significant problems in implementing a second task and having subjects engage in two tasks simultaneously. Subjects were not asked to think-aloud, a popular way of understanding users' cognitive activities, since they were already performing two tasks concurrently. To capture this type of data, a post-search questionnaire and exit interview were administered to each subject. The subjects appeared to be fairly well able to articulate their thoughts and perceptions in retrospect.

One interesting finding on the dual-task method is the importance of search time allotted to the subjects. In this experiment, the search time was restricted to 15 minutes. In their exit interviews only two subjects said they would have searched for up to an hour or longer beyond the allotted time; their search style differed from those of the other subjects in that both spent almost the entire search session reading one document. Both said they rarely bookmarked or printed out web pages as they read documents in entirety online. The other subjects tended to skim online documents and either bookmark or print out lengthy documents. Three subjects completed searches in less than the allotted 15 minutes. In a study of public library patrons, Slone (2002) found that the median length of online search sessions was 15 minutes. Subjects in the Fast and Campbell study (2004), who were limited to 5 minute searches, remarked that they would have done things differently if they had been given more time. In online search experiments, then, assigning too short a time for search sessions may distort subject responses or behavior.

Another timing-related issue with dual-task method implementation is determining an appropriate length of time between secondary task stimuli. The time intervals used in this study were relatively long (45-75 seconds) compared to other studies using dual-task methods (e.g., 5-10 seconds for Brünken et al., 5 seconds for Dennis et al.). The rationale for the longer intervals was to avoid having the secondary task become the dominant task because of its frequency. For example, in a 15-minute session with average interval length of 10 seconds, the secondary task stimulus is presented 90 times. But shorter interval times allow for the gathering of a larger number of data points, making possible some choice in the statistical analysis methods employed. Brunken et al. (2002) found statistically significant differences in response time for different modalities in their application of dual-task methodology.

Conclusion

The findings of this study indicate that dual-task methodology provides a useful technique for measuring mental effort during online searching. Two issues emerged as important decisions to be made when designing experimental studies using dual-task methodology. One is the length of the search session and the other one is the time interval between secondary task stimuli. Based on the results herein, 15 minutes appears to be a reasonable time for allowing subjects to engage in their usual search behavior. The time interval between stimuli used herein, 45-75 seconds, may have been too long, and so resulted in a relatively small number of measurements for both reaction time and miss frequency. In future research the interval should be shorter than the one used in this study, with the 5-10 seconds used in other dual-task experiments as a lower bound.

Future studies could investigate the use of more complex secondary tasks than the one used in this experiment. A relatively simple secondary task was chosen here because the researchers did not want the performance of the secondary task to take precedence over the primary task of searching, as one of the purposes of this research was to compare searching behavior on the web and in the library system. In the future, it may be desirable to design a study that does not compare the behaviors and focuses on measuring the degree of concentration or mental effort used in searching a particular IR system. Some possible secondary tasks for such a study are a memory retrieval task or a word production task.

Measuring a user's degree of concentration using a dual-task method also offers a way to assess task complexity. Rather than asking study subjects whether they perceive a search task to be difficult, complex, or familiar, task complexity may be assessed from the secondary task performance in a dual-task experiment. Multitasking information behavior is an important area of human information behavior (Spink, 2004), and a dual-task method could be used to measure and compare multitasking across different IR systems, or to compare the extent of the difference in complexity between multitasking and single-tasking. Such experimental results can then aid the design of systems supporting multitasking information behavior.

Acknowledgements

The authors thank Professor Karen Markey for providing insightful comments on developing questionnaires and interview questions, along with two anonymous reviewers for their suggestions on future studies.

Notes

[1] One of the nine subjects marked their confidence and difficulty levels for library searching as Not Applicable, so t-test was carried out with 7 degrees of freedom.

[2] Outliers outside the range of ±3 standard deviations around the original mean were excluded when computing the reported mean and standard deviation values

References

Borgman, C. L. (1996). Why are online catalogs still hard to use? *Journal of the American Society for Information Science, 47*(7), 493-503.

Britton, B. K., Muth, K.D., & Glynn, S. M. (1986). Effects of text organization on memory: Test of a cognitive effort hypothesis with limited exposure time. *Discourse Processes, 9*, 475-487.

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology, 49*(2), 109-119.

Cennamo, K. S. (1993). Learning from video: Factors influencing learners' preconceptions and invested mental effort. *Educational Technology Research and Development, 41*(3), 33-45.

Cooper, M. D. (2001). Usage patterns of a web-based library catalog. *Journal of the American Society for Information Science and Technology, 52* (2), 137–148.

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

Drabenstott, K. M. & Vizine-Goetz, D. (1994). *Using subject headings for online retrieval*. San Diego: Academic Press.

Eysenck, M. W. (2001). *Principles of cognitive psychology* (2nd Ed.). Psychology Press.

Fast, K.V. & Campbell, D. G. (2004). "I still like Google": University student perceptions of searching OPACs and the Web. *Proceedings of the 67th Annual Meeting of the American Society for Information Science & Technology, 41*

Glenberg, A. M. & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition, 15*(1), 84-93.

Grosjean, M., Rosenbaum, D. A., & Elsinger, C. (2001). Timing and reaction time. *Journal of Experimental Psychology: General, 130*(2), 256-272.

Iani, C., Gopher, D., & Lavie, P. (2004). Effects of task difficulty and invested mental effort on peripheral vasoconstriction. *Psychophysiology, 41*, 789-798.

Krendl, K. A. (1986). Media influence on learning: Examining the role of preconceptions. *Educational Communication & Technology Journal, 34*(4), 223–234.

Maki, R. H. & Berry, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10*(4), 663-679.

O'Donnel, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and performance: Cognitive processes and performance* (pp. 42.1-42.49). New York: Wiley.

Pashler, H. (2000). Task switching and multitask performance. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 277-309). Cambridge, MA: The MIT Press.

Pew Internet Project (2004). *The Internet and Daily Life*. Retrieved from http://www.pewinternet.org/pdfs/PIP_Internet_and_Daily_Life.pdf

Pew Internet Project. (2005). *Search engine users*. http://www.pewinternet.org/pdfs/PIP_Searchengine_users.pdf

Rieh, S. Y. (2005, January). *Is web searching easier than library searching?: Differential expectation, perceived difficulty, and effort*. Paper presented at the ALISE Annual Conference, Boston.

Salomon, G. (1981). Introducing AIME: The assessment of children's mental involvement with television. In K. Kelly and H. Gardner (Eds.), *New directions for child development: Viewing children through television* (pp. 89-102). San Francisco: Jossey-Bass.

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*(4), 647-658.

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

Slone, D. J. (2002). The influence of mental models and goals on search patterns during web interaction. *Journal of the American Society for Information Science and Technology, 53*(13), 1152-1169.

Spink, A. (2004). Multitasking information behavior and information task switching: An exploratory study. *Journal of Documentation, 60*(4), 336-351.

Spink, A., Jansen, B. J., Wolfram, D. & Saracevic, T. (2002). From E-Sex to E-Commerce: Web search changes. *IEEE Computer, 35*(3), 107-111.

Spink, A. & Jansen, B. J. (2004). *Web search: Public searching of the Web*. Boston: Kluwer Academic Publishers.

Sutcliffe, A.G., Ennis, M. & Watkinson, S. J. (2000). Empirical studies of end-user information searching. *Journal of the American Society for Information Science, 51*(13), 1211–1231.

Verwey, W. B. & Veltman, H. A. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. *Journal of Experimental Psychology: Applied, 2*(3), 270-285.