

Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

# Information Processing and Management

journal homepage: [www.elsevier.com/locate/infoproman](http://www.elsevier.com/locate/infoproman)

## Amount of invested mental effort (AIME) in online searching

Soo Young Rieh\*, Yong-Mi Kim, Karen Markey

School of Information, University of Michigan, North Quad, 105 S. State Street, Ann Arbor, MI 48109-1285, United States

### ARTICLE INFO

#### Article history:

Received 8 February 2012

Received in revised form 29 April 2012

Accepted 7 May 2012

Available online 3 July 2012

#### Keywords:

Amount of invested mental effort

Search effort

Search self-efficacy

Web searching behavior

Search system perception

### ABSTRACT

This research investigates how people's perceptions of information retrieval (IR) systems, their perceptions of search tasks, and their perceptions of self-efficacy influence the amount of invested mental effort (AIME) they put into using two different IR systems: a Web search engine and a library system. It also explores the impact of mental effort on an end user's search experience. To assess AIME in online searching, two experiments were conducted using these methods: Experiment 1 relied on self-reports and Experiment 2 employed the dual-task technique. In both experiments, data were collected through search transaction logs, a pre-search background questionnaire, a post-search questionnaire and an interview. Important findings are these: (1) subjects invested greater mental effort searching a library system than searching the Web; (2) subjects put little effort into Web searching because of their high sense of self-efficacy in their searching ability and their perception of the easiness of the Web; (3) subjects did not recognize that putting mental effort into searching was something needed to improve the search results; and (4) data collected from multiple sources proved to be effective for assessing mental effort in online searching.

© 2012 Elsevier Ltd. All rights reserved.

### 1. Introduction

It is well known that people conduct searches in brief sessions, entering short queries while rarely using advanced system features (Spink & Jansen, 2004). They also expect their searches to be immediate, instantaneous, and convenient (Markey, 2007). They think their searches are successful most of the time (Markey), feel confident about their searching abilities (Fallows, 2005), and are satisfied with the results (Markey). In fact, people's online searching behavior has hardly changed over the 25 years that online systems have been available to them (Markey). Despite the fact that Web searching is a daily activity for many people, they do not appear to be learning much about online searching.

The authors believe that people's simple searching behavior and their failure to learn more about online searching is related to the extent of mental effort they invest in online searching. Based on the theoretical framework of media perception and mental effort developed in the field of educational psychology (Cennamo, 1992–1993; Krendl, 1986; Salomon, 1984), this research posits that people's perception of information retrieval (IR) systems and their self-confidence in searching are closely related to the extent of mental effort they invest in searching. If people perceive that an IR system is easy to use, they are likely to put little mental effort into searching, whereas they would invest more mental effort if they believe the system is demanding and intended for serious use. On the other hand, people's self-confidence may also influence the extent of their mental effort in searching so that people should put more effort into searching and be more persistent at search tasks as they become more confident about their searching ability.

\* Corresponding author. Tel.: +1 734 647 8040; fax: +1 734 615 3587.

E-mail addresses: [rieh@umich.edu](mailto:rieh@umich.edu) (S.Y. Rieh), [kimym@umich.edu](mailto:kimym@umich.edu) (Y.-M. Kim), [ylime@umich.edu](mailto:ylime@umich.edu) (K. Markey).

This research examines the problem of simple online searching behavior by comparing the effort people exert when using two different information retrieval systems—a Web search engine and a library system. These two systems make for an interesting comparison as both systems are designed to support people in finding information, but people can perceive them to be two rather different systems. On one hand, students primarily use a university library system to find information for academic-related tasks. Since assignments and projects tend to be “imposed queries” (Gross, 1995), students may feel uncertainty (Kuhlthau, 2004) and even anxiety (Mellon, 1986) about library systems. On the other hand, people use the Web for a much broader range of purposes, for example, searching the Web to find people, news, information on a variety of topics, watching video, and listening to music. Therefore, people may feel that Web search engines are easy, fast, and convenient whereas library systems are complex, difficult, and demanding (Fast & Campbell, 2004). A comparison of online searching behavior in Web searching and library searching was also reported in detail in an earlier paper by the first author (Rieh, 2007). Further, people’s self-confidence about their ability to search library systems and Web search engines differs. According to a Pew Internet Report (Fallows, 2005), 92% of search engine searchers were confident in their searching abilities. In contrast, college students who participated in Fast and Campbell’s (2004) study said that they were less confident in their ability to search the library system effectively compared to the Web because of their lack of experience with the library system.

Building on Fast and Campbell’s results, the authors of this study claim that people’s differential perceptions of information systems and their self-efficacy in their searching abilities influence the amount of mental effort they put into searching, and, as a result, their differential mental effort impacts their search experience. To investigate this claim, our research enlists two experimental studies due to the challenges of measuring the extent of mental effort in the context of online searching. In Experiment 1, self-report measures are used to assess mental effort based on the assumption that investment of mental effort is an intentional process under control of the individual (Salomon & Leigh, 1984). In a follow-up study, Experiment 2 employs a dual-task method in which the subjects are engaged in two tasks simultaneously to measure the extent of mental effort invested into online searching (primary task) with respect to visual stimulus provided through a small window displayed on the computer screen (secondary task). The assumption of the dual-task method is that performance in one task diminishes in the presence of another task, due to limited cognitive capacity (Cennamo, 1993).

The two experimental studies share these research purposes: (1) to determine whether an examination of mental effort provides a useful conceptual framework for understanding people’s searching behavior; (2) to identify appropriate research methods that will allow researchers to measure the amount of mental effort that people exert in online searching activities. This paper addresses the following four research questions:

1. To what extent does people’s investment of mental effort in online searching differ when using a Web search engine and in an online library system?
2. How do people’s perceptions of these two systems and their level of self-efficacy affect their mental effort?
3. To what extent does differential investment of mental effort in the two systems influence user search experience?
4. Which data collection method (self reports vs. dual-task performance) provides a more reliable data set for measuring mental effort in online searching?

## 2. Related literature

### 2.1. User effort in search and retrieval

With a few exceptions, most studies in the field of online searching and information retrieval (IR) have investigated user effort implicitly rather than explicitly. Belkin and Muresan (2004) explicitly include user effort as one of the user-oriented evaluation measures for IR performance. They claim that degree of interaction is an important aspect of IR evaluation and identify user satisfaction, user effort, and efficiency as three major categories for assessing interaction process and outcome. They also propose that user effort could be measured through click counting, iterations, scrolling, documents seen, viewed or read. Gerwe and Viles (2000) focus on studying user effort in query construction. Their study enlists number of search terms and advanced search features as measures of user effort.

User effort could be a key to explaining the gap between interaction-oriented evaluation and algorithm-oriented evaluation in IR. Because previous studies suggest that system-oriented performance measures and user evaluation do not always give the same results (e.g., Turpin & Hersh, 2001), improvements in an information system’s effectiveness do not directly translate into benefits for users. Turpin and Scholer (2006) examine the problem using two different tasks: a precision-oriented task requiring users to find one document and a recall-based task that measured the number of relevant documents users found. The results of their study show that there is no significant relationship between system effectiveness and the precision-based task; however, a weak relationship was presented between the performance measure and the simple recall-based task. Another study by Allan, Carterette, and Lewis (2005) addresses the question of “how much better do search systems need to be for people to appreciate the improvement” (p. 433). Their results reveal that users tend to gain only little unless they can recognize that system accuracy has improved to a great extent.

To study this problem, Smith and Kantor (2008) investigated users’ adaptive search behaviors. In their experiments, they manipulated the search results, presenting search results in one of three ways to subjects: (1) in the “consistently-low-rankings condition,” Google’s 300th and lower-ranked retrievals are displayed, (2) in the “inconsistently-low-rankings condi-

tion,” Google’s 1st to 320th ranked retrievals are displayed inconsistently (i.e., in mixed up order), and (3) standard results from Google, with the number 1-ranked item first and subsequent items returned in ranked order. They found that subjects who used one of the degraded systems were as successful finding relevant retrievals as those using the standard system. To achieve this success, the searchers in their study modified their searching behavior in ways that depended on the characteristics of the failure.

These previous studies indicate that people put effort into retrieval processes. When the performance of a system is “good enough,” in general, people accept retrieval results without being aware of the incremental improvement in system performance (Allan et al., 2005). When the system performs badly, users rise to the occasion, adapting themselves to degraded systems to effect results that satisfy their information needs (Smith & Kantor, 2008).

## 2.2. Amount of invested mental effort (AIME)

Salomon (1981) introduced Amount of invested mental effort (AIME) as a way to explain apparently contradictory findings on the relationship of media, mental effort, and learning. Because AIME is defined as “the number of non-automatic mental elaborations applied to a unit of material” (p. 92), it reflects a voluntary allocation of effort that can be reported by the individual. Salomon explains that people 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 people perceive the encountered material to be easy, warranting little investment of mental effort on their part.

A number of variables can affect the amount of invested mental effort (AIME). Salomon (1983, 1984) identifies two primary factors: Perceived Demand Characteristics (PDCs) of the stimulus, task, or context, and Perceived Self-Efficacy (PSE). With regard to PDC, people may invest more effort in a medium when they perceive it to be “serious” and “demanding” (Salomon, 1983, p. 47). For example, in a study comparing learning from print and television, Salomon (1984) found that young people learned less from television because they perceived television to be “easy” and print “hard.” His conclusion was that perceiving a medium as being easy, as in the case of television, led to people investing less mental effort in the activity, resulting in less learning.

Salomon and Leigh (1984) note that one of the easiest ways to change PDC is to change people’s perception of the task. For example, before being shown a videotape, half of the children were instructed to watch or read the story “for fun” (low PDC) while the other half were instructed to watch or read the story to “see how much you can learn from it” (high PDC). According to Salomon and Leigh, the former condition represents typical television viewing and the latter represents a condition in which viewers perceive TV more seriously. They found that people exerted more mental effort in both TV and print stories under high PDC conditions. For example, in experiments with undergraduate students, Kunkel and Kovaric (1983) found that subjects reported investing more mental effort processing TV programs when the researchers told subjects that the programs were broadcast on PBS than when they told them the programs were broadcast on the commercial networks. Subjects associated PBS programming with “harder” educational viewing and commercial network programming with “easier” entertainment viewing.

A second factor affecting AIME is the perception pertaining to people’s Perceived Self-Efficacy (PSE) in processing materials. Self-efficacy, introduced by Bandura (1982), has been employed to explain people’s performance differences in settings such as the classroom, work organization, and athletic performance. According to Bandura (1997), “Perceived self-efficacy refers to beliefs in one’s capabilities to organize and execute courses of action required to produce given attainments” (p. 3). Self-efficacy is usually measured with questions that ask people their degree of self-confidence regarding their ability to perform specific tasks. People with high levels of self-efficacy are more likely to persist and invest mental effort at a task when individual abilities and task difficulty are of similar levels: “When successes are hard to come by, individuals of high efficacy are persisters and those of low efficacy are rapid quitters” (Bandura, 1997, p. 216).

Clark (2001) has found an “inverted U” relationship between self-efficacy and amount of mental effort invested. According to Clark, mental effort slows, then stops, at either exceptionally low or high self-efficacy levels. When tasks are perceived as impossible, people are likely to avoid the goal at hand. The more novel and difficult the goal is perceived to be, the more challenge is expected in the task, and “the more effort we expend until the novelty grows beyond an efficacy threshold” (Clark, 2001, p. 272). The more familiar the goal and the more knowledge and skill have been gained in the pursuit of the same goals, the less effort is invested. Eventually, self-efficacy is reached to its maximum and an “overconfident default” can occur. The crux of the issue seems to be how to draw the greatest amount of mental effort for solving novel tasks without exceeding the efficacy threshold.

## 2.3. Methods of assessing mental effort

Various methods and instruments for measuring mental effort have been proposed and used in the fields of educational psychology, experimental psychology, and human factors engineering. These methods fall under three categories: (1) self-report measures, (2) cognitive workload measures, and (3) physiological measures (Gimino, 2000). Self-report measures assume that subjects can accurately assess and report their cognitive processes because “the investment of effort is a voluntary process that is under the control of the individual and as such is available for introspection” (Cennamo, 1993, p. 36). Cognitive load theory (Sweller & Chandler, 1991) is the basis for cognitive workload measures, and dual-task methodology is one

of the methods used for this purpose. In dual-task methodology, subjects are usually assigned one primary task and one secondary task, where the primary task may be reading a passage, viewing a television program, or doing a Web search, and the secondary task may be responding to a tone or a flash of light, or performing mental arithmetic (Kim & Rieh, 2005). The assumption is 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, 1993, p. 37). The extent to which a person’s performance on a secondary task is reduced (e.g., slower response time, missing a cue) is a measure of the primary task’s mental effort requirement. Lastly, physiological measures, such as heart rates (Veltman & Gaillard, 1998) and brain activity measurements (Gevins et al., 1998) 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).

A few studies have used dual-task measures to test the reliability and validity of self-report measures (Gimino, 2000) and to explore the relationship of task difficulty and self-reported mental effort (Flad, 2002). Gimino compared self-reports of Salomon’s AIME and Paas’s (1992) mental effort scale against dual-task performance results of 207 high school students. Subjects had to complete sets of mental addition problems varying in novelty while interrupted by a visual stimulus. Gimino found a curvilinear relationship between cognitive workload and task novelty, supporting Clark’s (2001) “Inverted U” relationship between self-efficacy and mental effort. In Flad’s study, subjects had to respond to a visual stimulus while trying to memorize and enter sets of digits into a computer. Flad found that for very difficult problems, subjects reported using more mental effort but the dual-task measures indicated that the subjects had actually decreased the amount of mental effort used.

A very small number of studies measure mental effort in online searching. Dennis, Bruza, and McArthur (2002) enlisted the dual-task method to quantify user efforts in various search states. Subjects performed searches in one of three systems, Yahoo!, Google, and a system developed by the researchers. While performing searches, subjects had to respond to an auditory stimulus. The results were mixed: while there were significant differences between the search systems in secondary task performance when a search time limit was imposed, no such differences were found when searches were self-terminated.

### 3. Model of invested mental effort in online searching

The model presented in Fig. 1 shows the primary variables to be examined in this study. This study examines two kinds of Perceived Demand Characteristics (PDC) in online searching: perceptions of information retrieval (IR) systems and perceptions of search tasks. Both kinds of perceptions are positively related to amount of invested mental effort (AIME) because people are likely to invest more effort in a system or in a search task when they perceive it to be serious and demanding.

Bandura’s (1997) theory supports only a positive relationship between self-efficacy and mental effort. The implication of Bandura’s theory to this study is that a person is likely to invest more mental effort if he or she is very confident about searching in general and he or she likes challenging and novel search tasks. Based on Salomon’s (1984) AIME (amount of invested mental effort) theory, however, the idea underlying the model is that the PDC (Perceived Demand Characteristics) of the task and the PDC of the system are related to PSE (Perceived Self-Efficacy), and both affect the AIME that people exert. Unlike PDC which is positively related to AIME, PSE can either positively or negatively influence AIME. When PSE is high, AIME may be higher for more demanding tasks or systems and lower for tasks or systems that people perceive as being easy.

It should be noted that PSE and PDC are negatively related to each other in the model. This is because if a person perceives an IR system or a search task to be difficult or demanding, she/he would be likely to feel low self-efficacy and vice versa. The model also presents that the extent of AIME is somehow related to user search experience. However, how exactly AIME is related to user search experience is as yet to be determined. AIME may affect user search experience positively at the conclusion of successful searches in which people invested considerable mental effort. On the other hand, AIME may have a negative influence on user search experience, for example, when people express frustration and anxiety as a result of putting so much effort into searching.

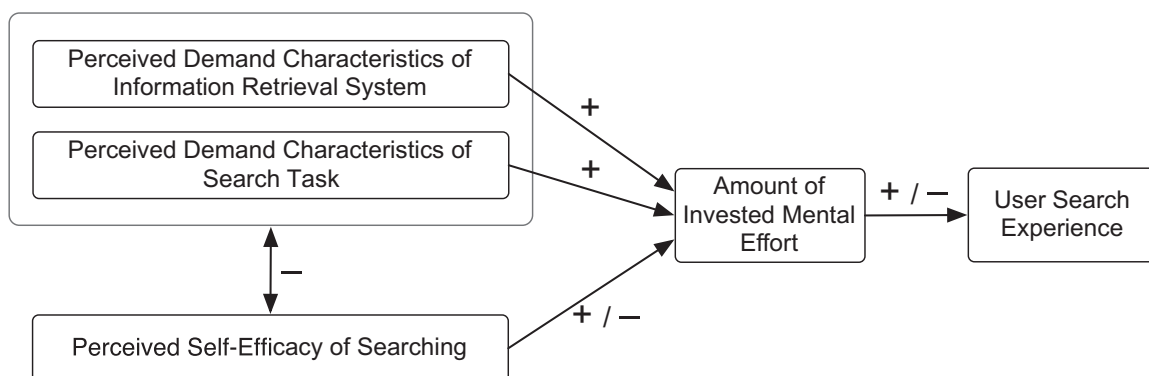


Fig. 1. Amount of invested mental effort (AIME) in online searching.

#### 4. Experiment 1: self-reports based mental effort with two search tasks

A major difference between the two experiments conducted in this research was the method of assessing mental effort. Experiment 1 was designed to measure mental effort primarily based on self-reports. In Experiment 1, the PDC was examined in two ways—perception of information retrieval (IR) systems and perception of search tasks. The data were collected through pre-search background questionnaires, think-aloud protocols, search transaction logs, post-search questionnaires, and post-task interviews.

##### 4.1. Method

Fifteen undergraduate students at the University of Michigan (U-M) were recruited for Experiment 1. The subjects volunteered to participate in the study by responding to flyers that the researchers posted in university buildings. Flyers requested the participation of students who had been given a term paper, course assignment, or project to complete and required the use of library resources. Volunteers were five freshmen, three sophomores, one junior, and six seniors. Ten were male and five were female. Their major areas were rather diverse: biology, sociology, psychology, French, nursing, business, and political science. They received monetary compensation for their participation which took about 1 hour.

Subjects conducted four searches in total. The researcher gave them two tasks to do in two systems. For the library system search condition, subjects searched the U-M Library's homepage (<http://www.lib.umich.edu>). This library page included links to diverse library resources such as the Mirlyn Classic library catalog, abstracting and indexing databases, electronic journals and newspapers, library services and other information. For the Web search condition, subjects were allowed to start in any Web search engine site that they use most frequently.

The two tasks used in this study represent typical search problems for each system. Searching for an academic purpose is the most popular activity on a library system (Rieh, Hilligoss, & Yang, 2007). Searching for product information on the Web is a common task for Internet users to do (Jansen, 2010). We designed the Research Task in generic ways so that students can adapt their own research topic. It was because we could assume that students are fairly familiar with this type of search task. The Product Task was less generic because there might be high level of variance in product search among college students depending on their familiarity or experience with this topic. The researchers gave subjects two tasks in a written form as follows:

1. I would like you to do a search for your term paper topic. If you've already done some research, just pick up from where you left off. Don't repeat searches you've done in the past [Research Task].
2. You want to buy a new digital camera and you need to decide which model you would buy. You would like to read some reviews about digital camera to help you make a decision [Product Task].

The search conditions were randomized in a way that allowed each subject to pick the order of four searches. Upon arriving at the lab, subjects signed a consent form and completed a pre-search background questionnaire. This questionnaire asked for demographic information, search experience, and their perception of the "easy-ness" of searching the Web and library systems, and their self-confidence using online systems. The subjects were also asked how demanding they thought each search task (Research Task and Product Task) would be. While conducting searches, subjects were asked to think-aloud and tell what they were doing and why. Camtasia, software for capturing screen activities, was used to save search logs and an audio tape recorder recorded subjects' think-aloud verbal reports. Every time subjects completed a search, they filled out a post-search questionnaire that collected information about their difficulty searching, the effort they put into searching, and other learning and searching experiences. Each time subjects finished two searches for one task, a post-task interview was conducted to identify their searching experiences using the library system and the Web for that particular task.

For all 15 subjects, audiotapes of the think-aloud and interview recordings were transcribed and content-analyzed. The search logs saved as Camtasia files were played back and manually coded based on the coding scheme which was developed to measure the users' mental effort in terms of time taken, number of clicks, and number of actions taken by the subjects. The coding scheme drawn from the log analysis include:

- Time taken to complete a search session.
- Time taken to formulate the first query.
- Time taken to view the search results.
- Time taken to read full-text documents retrieved.
- Number of clicks per search session.
- Number of search results selected per session.
- Number of full-text documents read per session.
- Number of words entered in the first query.
- Number of words entered in average per query.
- Number of query reformulations during session.
- Number of advanced features used during session.

## 4.2. Results

### 4.2.1. Analysis of self-reports

The results from the self-reported questionnaires indicated that subjects had significantly different levels of perceived difficulty for Web searching and library system searching. The responses from the pre-search background questionnaire showed that subjects perceived that library system searching would be significantly more difficult ( $M = 4.67$ ,  $SD = 2.16$ ) than Web searching ( $M = 1.67$ ,  $SD = 2.09$ ),  $t(14) = 4.3$ ,  $p < .01$  on a scale of 0 (not difficult) and 10 (very difficult). In the pre-search questionnaire, subjects also responded that they were significantly more confident with Web searching ( $M = 8.4$ ,  $SD = 1.24$ ) than library system searching ( $M = 5.13$ ,  $SD = 2.87$ ),  $t(14) = 4.23$ ,  $p < .01$  on a scale of 0 (not confident) and 10 (very confident).

The responses from the post-search questionnaire showed that subjects' initial perceptions—Web searching was easier than library system searching—had not been changed after they searched both systems; however, when comparing the perceived difficulty between before and after searching, it was found that Web searching turned out to be more difficult than subjects originally perceived. When they completed Web searching for the Research Task, they reported on a scale of 0 (not difficult) and 10 (very difficult) that it was difficult to choose right keywords to present their ideas of information (4.33), to formulate a search query to enter into the system (3.53), to understand search results (4.53), and to make decisions about which information (papers and Web pages) were useful (4.67) (see Table 1). The only figure that was comparable to the 1.67 in the pre-search questionnaire about Web searching was the difficulty starting the search (2.2). In the case of library system searching, their initial perception of library searching (4.67) was either slightly increased or remained in the same level. When subjects engaged in the Research Task using the library system, they reported somewhat higher levels of difficulty connected with choosing right keywords to present their ideas of information (5.66) and formulating search query to enter the system (5.53) compared to their perceptions prior to searching. The difficulty perceived in understanding search results (4.47) and making decisions about which information (papers and Web pages) were useful (4.87) remained on the same level as their initial perception.

Even before conducting searches, the subjects perceived that the Research Task would be significantly more demanding ( $M = 5.33$ ,  $SD = 2.77$ ) than the Product Task ( $M = 2.93$ ,  $SD = 2.87$ ),  $t(14) = 2.8$ ,  $p < .05$  on a scale of 0 (not demanding) and 10 (very demanding). This differential perception of search tasks directly influenced the amount of invested mental effort (AIME) that subjects exerted. As shown in Table 2, the difference between AIME in Web searching and library system searching was much greater for the Product Task than for the Research Task. For Product Task, subjects responded that they invested more effort (Web  $M = 4.53$ , Lib  $M = 6.8$ ), put more thoughts (Web  $M = 3.87$ , Lib  $M = 6.73$ ), and gave greater concentration (Web  $M = 5.27$ , Lib  $M = 7.4$ ) to library system searching than to Web searching. When subjects engaged in the Research Task, however, they put more or less same level of effort (Web  $M = 6.06$ , Lib  $M = 6.67$ ), put almost the same amount of thoughts (Web  $M = 6.33$ , Lib  $M = 6.6$ ), and gave similar amount of concentration (Web  $M = 6.73$ , Lib  $M = 7.13$ ) to both library system searching and Web searching.

These findings indicate that when people perceive the search task to be demanding, their perceptions of IR system may not matter. Given the perception of the search task to be demanding, subjects may feel that they have to invest certain amount of mental effort no matter what kind of information system they are using. On the other hand, when subjects deal with an easy search task such as product information, the kind of system seems to make a difference in the amount of invested mental effort (AIME).

### 4.2.2. Analysis of search transaction logs

This section examines differences in the investment of mental effort between the library system and the Web based on the analyses of search transaction logs. Apparently there were discrepancies between how subjects perceived difficulty in searching and how they actually perform searches. Although subjects reported that they felt greater difficulty in formulating the search queries they entered into the library system ( $M = 5.53$ ) than into the Web search engine ( $M = 3.53$ ), the transaction logs reveal that they actually spent less time formulating their first query in the library system search (17 seconds) than on the Web search (22 s) for Research Task. Subjects spent even less time in formulating the first query for the Product Task than for the Research Task. It took them only 9 s in the library system and 11 s on the Web even though the subjects rated formulating queries in the library system ( $M = 5.13$ ) as being more difficult compared to the Web ( $M = 1.13$ ).

The average number of query words could be an indicator of mental effort. The results show that AIME in searching and the number of words entered are not necessarily related directly. Unlike subjects' perceptions that they put more effort into library system searching in general, they entered slightly longer queries into the Web than into the library system. Table 3 shows that for the Product Task subjects entered 2.8 words into the Web search engines and 2.36 words into the library system. They also entered longer queries for the Research Task than the Product Task, entering 3.46 words into the Web search engines and 3.05 words into the library system.

It was believed that the following measures of search activities would be particularly useful to assess mental effort: number of clicks, number of query reformulations, and number of advanced features used during the session. The results of transaction logs analyses seemed to be consistent with how subjects responded in the post-search questionnaires. Apparently, subjects put more effort into searching the library system than they did for Web searching. For both Product and Research Tasks, subjects clicked more links in the library system (47.57 clicks for Product Task and 41.67 clicks for Research Task) than

**Table 1**  
Perceived search difficulty from Experiment 1.

Questions	Product Task		Research Task	
	Web search mean (SD)	Lib. search mean (SD)	Web search mean (SD)	Lib. search mean (SD)
Difficult to start the search (P: <sup>**</sup> , R: <sup>*</sup> )	1.06 (1.62)	6.0 (3.16)	2.2 (2.30)	4.47 (2.85)
Difficult to choose keywords to present your idea (P: <sup>**</sup> )	1.07 (1.27)	5 (2.78)	4.33 (2.87)	5.66 (3.04)
Difficult to formulate search query in the system (P: <sup>**</sup> , R: <sup>*</sup> )	1.13 (1.35)	5.13 (2.47)	3.53 (3.00)	5.53 (2.85)
Difficult to understand search results (P: <sup>**</sup> )	2.33 (2.89)	5.6 (2.44)	4.53 (2.92)	4.47 (3.16)
Difficult to make decisions about usefulness of information	3.26 (3.08)	4.8 (3.21)	4.67 (2.63)	4.87 (2.90)

Note: A scale of 0 (not difficult) – 10 (very difficult).

P refers to Product Task, R refers to Research Task.

\* Indicates  $p < 0.05$ .

\*\* Indicates  $p < 0.01$ .

**Table 2**  
Perceived mental effort from Experiment 1.

Questions	Product Task		Research Task	
	Web search mean (SD)	Lib. search mean (SD)	Web search mean (SD)	Lib. search mean (SD)
Effort invested into searching (P: <sup>*</sup> )	4.53 (2.80)	6.8 (1.97)	6.06 (2.15)	6.67 (1.84)
Thoughts put into searching (P: <sup>**</sup> )	3.87 (2.42)	6.73 (2.15)	6.33 (1.99)	6.6 (1.99)
Concentration during the search (P: <sup>**</sup> )	5.27 (2.54)	7.4 (1.68)	6.73 (1.75)	7.13 (2.03)
Usefulness of search results (P: <sup>**</sup> )	7.53 (2.69)	2.67 (2.99)	5.67 (2.63)	4.93 (3.49)
Importance of search (P: <sup>**</sup> )	7.86 (2.87)	2.0 (1.96)	5.07 (2.71)	4.6 (3.11)

Note: A scale of 0 (no effort, no thinking, not concentrated) – 10 (a great deal of effort, a great deal of thinking, very concentrated).

P refers to Product Task, R refers to Research Task.

\* Indicates  $p < 0.05$ .

\*\* Indicates  $p < 0.01$ .

**Table 3**  
Search transaction log analysis from Experiment 1.

Search transaction logs	Product Task		Research Task	
	Web search mean	Lib. search mean	Web search mean	Lib. search mean
Time taken to complete a search session (min/s)	12:59	12:16	12:37	12:53
Time taken to formulate the first query (min/s)	0:11	0:09	0:22	0:17
Time taken to view the search results (min/s)	2:29	5:17	3:41	5:08
Time taken to read documents (min/s)	8:40	3:12	6:21	3:11
Number of clicks per search session	37.13	47.57	35.8	41.67
Number of search results selected per session	6.53	7.57	10.73	7.27
Number of words entered in the first query	2.8	2.18	2.8	2.37
Number of words entered in average per query	2.8	2.36	3.46	3.05
Number of query reformulations during session	1.53	5.79	2.93	5.2
Number of advanced features used during session	0.47	3.21	1.07	3.87
Number of documents viewed per session	16.87	7.5	13.67	7.47

the Web (37.13 clicks for Product Task and 35.8 clicks for Research Task) as presented in Table 3. They also reformulated their queries more often in the library system: in the case of the Product Task, they reformulated queries 5.79 times, and for the Research Task, they reformulated queries 5.2 times. While using the Web search engines, they reformulated queries 1.53 times for the Product Task and 2.93 times for the Research Task respectively. In addition, subjects used advanced features such as Boolean operators, plus sign, minus sign, or quotation marks more often when they used the library system (3.21 times for the Product Task and 3.87 times for the Research Task) than on the Web (0.47 times for the Product Task and 1.07 times for the Research Task).

#### 4.2.3. Analysis of think-aloud and interviews

Content analyses of subjects' comments from think-aloud during the searches and post-task interviews revealed a number of interesting findings. The results could fall into the following four themes: perceptions of IR systems, self-efficacy, mental effort, and user search experience.



*4.2.3.1. Perceptions of IR systems.* The results from the analyses of interview and think-aloud data were consistent with the responses from the questionnaires. Subjects indeed were able to describe different perceptions about the Web search engines and the library system. Most subjects perceived that the Web search engines were quick, easy, not at all difficult, and self-explanatory whereas the library system was difficult to use, difficult to decide how to start a search, harder to find things, and not easy to use. Here are some examples how the subjects described the characteristics of Web searching.

- Well, the Web search wasn't hard at all because I knew exactly where I was going. I was really familiar with it. (S05)
- I knew in general what I was looking for and I could kind of narrow it down from broad search topics a little bit more easily on the Web than the library search. (S06)
- It gives you places to throw your keywords in. If you know what you're talking about, you just put the keyword. (S08)

It was noted that the subjects often talked about their perceptions of IR systems in terms of differential expectations of each system rather than in terms of their actual search experiences. To many subjects, the library system was all about classes and school. S03, for instance, described, "In the library it was nearly impossible because this seems entirely geared toward research that is more academic like a professor is going to do a paper on digital camera, so nothing showed up." S04 said, "I use the Google search which is basically using the Web, the library is basically going to be scientific journals and I was looking more for personal opinions and feedback which is harder to get [in the library]." S12 also said that "I just imagine the library Web site to be for homework, papers, and projects."

Even though the library system has provided electronic resources for over a decade, subjects have strong preconceptions that the library system is all about books and the Web is about online information resources. For instance, S08 made the following comments before searching the library system, "Okay first thing I'll try to search in the library is Mirlyn [the University of Michigan Library's online catalog system] because they have good databases there. Problem with Mirlyn is obviously I really can't get any information right now. I would have to go to the library to pick up the books so I'm not really sure how helpful this is going to be for the final project." S15 said, "But I don't know how helpful that will be because I just think they're going to give me the books but I'm going to try it anyway."

This differential expectation was also found in regard to the extent of trust subjects had for the Web search engines and the library system. Overall, subjects were aware of the potential problems of information credibility in Web sites, and they took precautions about using Web information for the Research Task. S07, for instance, said that "The thing about just looking at Web sites in general, I'm not sure how trustable [sic] they are." S14 also stated, "We don't know if it's a good and reliable source." The subjects often talked about Web information credibility with respect to the type of search task. They were more concerned with credibility issues when they looked for information for the Research Task. S02 mentioned that "I never use Yahoo for a paper just because you find so many unreliable Web sites that don't have any citations." S04's comments revealed his trust in library resources, "Plus the library gives you the advantage that you know that it's actually a good source. And the person who wrote the book knew what they were talking about." S14 even named specific databases that he could trust, "I like ProQuest because they are academic journals so it's trustable [sic] and definitely something I want for my class. . ." Subjects seemed to accept the product information they found on the Web more easily than Web-based information for their research; however, they were still aware of the problem of credibility with Web-based product information. S03 made strong comments about the challenges of Web searching, "Possibly sometimes when you search on Google, a lot of the links are junk, like random pages that aren't related or are just something made by some person what says a whole Web site that made as an individual about this camera is great. And that's okay but it's not from an authoritative source. Like it's hard sometimes to weed through those links on Google search."

*4.2.3.2. Self-efficacy.* Across questionnaire responses and interviews, subjects reported that they had different levels of self-efficacy between the library system and the Web search engine. When asked what would account for a searcher's failure to successful searches in the library system, S03 replied, "Not knowing how to use the library. Like whether to use Mirlyn [the U-M Library's online catalog system] or any of the various links that are available there. They're very inept at using it. I feel only a little bit of confidence in it." It was found that familiarity and experience was the primary factor that influenced the level of self-efficacy most directly. S05 said that he used the U-M Library system only once or twice before, so he was not familiar with it whereas the Web search engine was something he was "really familiar with how to use it." S06 also stated that the reason that he did not feel being capable of using the library system was "just for not being as familiar."

The lack of confidence in searching often caused the feeling of being lost in the library system. During the think-aloud, many subjects searching the library system used the phrases such as "I don't know what's going on" (S15) or "I don't know where I am right now" (S13). Note how S15 talked about his feelings of being lost in this quote:

The first one seems to be way too specific for my topic. The next book however might actually be there. It doesn't seem as helpful. . . I don't know if I can get back to the main menu. I guess databases would be another thing I can look into. I don't know what that is exactly. Newspapers. . . okay that might be helpful to me but I don't know if I can search through this or if the search actually encompassed that before. (S15)

Subjects' initial assessment of their self-efficacy did not always remain consistent. Their initial perceptions of difficulty of using the system could change as they learned more about the system as a result of using it. The following quote is noteworthy in this regard:

So I would probably go back to the library because now I'm more familiar with how to use the system. It seems like it gives you better results. I thought that I got the results a lot quicker. I knew how to start with the Web better, but after I got used to it a little bit the results from the library were a lot better. (S05)

**4.2.3.3. Mental effort.** The analyses of think-aloud and interview transcripts indicated that subjects did not talk about the extent of effort they put into searching unless they were asked about it explicitly. It appeared that mental effort was not something people could easily express on their own especially during search sessions; however, during the interviews, subjects were able to identify various aspects of mental effort they put into searching. To our subjects, effort meant “not giving up and being persistent” (S02), “more patience and more willingness to explore it” (S03), and “looking all over the Web site trying to search” (S12). S03 believed that the level of effort put into searching was directly related to the success of searching saying that people's searches would be likely successful if “they've taken the time to click on each one of these and figure out what it is and go into the subtext of it and kind of feel that they're okay with it. Like it doesn't come down on them.” S02, on the other hand, described searching effort with respect to not being easily frustrated in the process. She specifically mentioned that people needed to take the following actions to make their search successful:

You have to find the right keywords and it's harder on the library. Also once you get to the library page sometimes you can find articles and you know they exist but it's really hard to find exactly where on the Web site you can find the actual text to the article... I think some people just get frustrated and want the article to be right in front of them as soon as they find it exists so they just give up. And some people only search for books. I notice that they don't really know about the electronic resources page which is usually the most helpful for me. (S2)

**4.2.3.4. User search experience.** Subjects' differential expectations toward systems affected the ways that they understood their search experiences. For instance, when subjects were asked why a person would not be able to find information using Web search engines, they responded that the Web must have that information and the problem was with the person who was not able to find the information on the Web. When the same question was asked with respect to the library system, subjects assumed that the library might not have the information. Because most subjects believed that the Web searching was going to be easy to use, they tended to blame themselves when they could not find information on the Web. On the contrary, when they struggled to find information in the library system, they blamed the library system—both its content and the design of its interface. Regarding a Web search engine, S11 said, “...just because I wasn't able to find everything. It's very possible it's all there I just wasn't able to connect to them.” However, when the search was not successful in the library system, her reaction was different: “Well first and foremost, the information is just not there, very possibly, I don't know if they have it” (S11). Another subject made comments that “I would be more prone to blame the Mirlyn [U-M Library online catalog]... because the format for it is just not a very good interface for someone who is trying to search for something specific” (S06).

It was found that subjects did not solve the complexity of interface in the library system by simply putting more effort into searching. Instead, they expressed affective feelings using terms such as “complicated,” “overwhelming,” “confusing,” “hard to understand,” and “difficult to use.” In contrast, their reaction to Web searching was consistently positive saying the Web search engine was “easy” and “familiar.” They were “willing to explore” and “enthusiastic” when using Web search engines. One subject explained the distinction between library system and Web search experiences as follows, “There's one [Web] search engine that will search everything and most people are familiar with it. Compared to the library one where they have tons of networked electronic resources that are there for different things. It's not just one encompassing it all. And I think that's where it makes it a lot easier on the Web” (S09).

## 5. Experiment 2: dual-task based mental effort

Experiment 2 was designed to measure mental effort using the dual-task method in which two tasks were performed concurrently. The overview of the dual-task method and a number of factors affecting dual-task performance were discussed in detail in an earlier paper by the authors (Kim & Rieh, 2005). The same data collection instruments (a pre-search background questionnaire, a post-search questionnaire, and a post-task interview) from Experiment 1 were re-used in Experiment 2. However, unlike Experiment 1, subjects were not asked to do think-aloud during their searches because they were already engaged in two tasks. Another difference was that Experiment 2 used only the Research Task in which the subjects were asked to do a search for their term paper topic, not the Product Task.

### 5.1. Method

Each subject performed two searches, one in the U-M Library system, and the other in a Web search engine of her/his choice. For “dual” tasks, the primary task was searching in either system (library or Web) and the secondary task was responding to a visual stimulus. This visual stimulus was provided through a small window displayed on the upper right area of the computer screen, to the right of the browser window. This window changed color, either from gray to red or red to gray, every 45–75 seconds. Subjects had to press the escape key on the keyboard as soon as they noticed the color change.

The color change time and the time at which the escape key was pressed were logged, and reaction time (RT) was computed as the difference between the two times. In addition to the dual-task searches, a single-task condition was included in the experiments to obtain baseline measures of each subject's reaction time. In the single-task condition, a blank browser window along with the small color window was displayed.

Study subjects carried out their searches using the Firefox browser running on an Apple G4 iBook with 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. A Java 1.4.2 applet provided the visual stimulus for the secondary task while a program called Global Hotkey associated the action of pressing the escape key with the resulting reaction of the system time being written to a data file.

Nine undergraduate students at the U-M participated in Experiment 2. Participants were recruited in two ways: by posting the experiment descriptions to the University website of paid experiments and sending out emails to class mailing lists in three U-M academic units (School of Information, the Department of Psychology, and the Department of Political Science). Volunteers were five freshmen, one sophomore, and three seniors, and six were male and three were female. Academic majors were computer science, economics, mathematics, psychology, music, biology, biochemistry, and two undecided.

The procedures in Experiment 2 were as follows. Subjects signed a consent form and completed a pre-search background questionnaire prior to starting their searches. While they searched the two systems, the experimenter observed subjects' searching behavior and took notes. After completing each search, subjects completed a post-search questionnaire that included questions on their difficulty in searching and the amount of mental effort invested in searching. At the end of all searches, an exit interview was conducted regarding their search experiences in using Web search engines and the library system. The entire interviews were recorded, transcribed and analyzed.

## 5.2. Results

### 5.2.1. Analysis of self-reports

Prior to carrying out their searches, subjects perceived that searching in the library system ( $M = 4.38$ ,  $SD = 2.45$ ) would be significantly more difficult than searching on the Web ( $M = 1.25$ ,  $SD = 0.89$ ) according to a two-tailed paired  $t$ -test,  $t(7) = 3.28$ ,  $p < 0.05$ , on a scale of 0 (not difficult) to 10 (very difficult). They were marginally more confident with Web searching ( $M = 6.5$ ,  $SD = 3.25$ ) than library system searching ( $M = 4.38$ ,  $SD = 2.26$ ) according to a paired  $t$ -test,  $t(7) = 2.15$ ,  $p = 0.07$  on a scale of 0 (not confident) and 10 (very confident).

Once subjects conducted searches, as presented in Table 4, they found it significantly ( $t(8) = 3.96$ ,  $p < 0.01$ ) more difficult to start their search in the library system ( $M = 5.0$ ,  $SD = 2.74$ ) than on the Web ( $M = 1.44$ ,  $SD = 1.94$ ) and significantly ( $t(8) = 2.80$ ,  $p < 0.05$ ) more difficult to choose keywords in the library system ( $M = 6.22$ ,  $SD = 2.33$ ) than on the Web ( $M = 3.33$ ,  $SD = 3.28$ ), on a scale of 0 (not difficult) and 10 (very difficult). Subjects also reported that they invested more effort (Lib  $M = 6.22$ , Web  $M = 4.56$ ) and concentrated more (Lib  $M = 8.0$ , Web  $M = 6.44$ ) into searching the library system than the Web, as shown in Table 5. No significant difference was found in the two systems in terms of the difficulty of formulating search queries, understanding search results, and making decisions on the usefulness of information.

### 5.2.2. Analysis of dual-task data

Subjects were presented with a visual stimulus 9–15 times during the task, the number of times being more or less depending on the duration of their searches. The stimulus appeared in a small window in the screen. Three measures of dual-task performance were used: Reaction Time (RT), Miss Frequency, and False Alarm. Reaction time is the amount of time it took a subject to respond to the visual stimulus, measured in milliseconds. Miss Frequency referred to the number of times a subject failed to respond to the visual stimulus. False Alarms recorded events in which a subject responded even though a visual stimulus was absent.

The analysis of reaction times of the three conditions—Single-Task (secondary task only), Library Search Dual-Task, and Web Search Dual-Task—showed that the Single-Task condition had the lowest mean reaction time ( $M = 0.6$  s,  $SD = 0.11$ ), followed by the Web Search Dual-Task condition (3.31 s,  $SD = 2.06$ ) and the Library Search Dual-Task condition ( $M = 4.84$  s,  $SD = 6.52$ ). Differences in reaction times between the Single-Task condition and the Web Search Dual-Task condition

**Table 4**  
Perceived search difficulty from Experiment 2.

Questions	Research Task	
	Web search mean (SD)	Lib. search mean (SD)
Difficult to start the search**	1.44 (1.94)	5.0 (2.74)
Difficult to choose keywords to present your idea*	3.33 (3.28)	6.22 (2.33)
Difficult to formulate search query in the system	3.56 (3.40)	5.0 (2.00)
Difficult to understand search results	2.67 (2.00)	4.33 (2.60)
Difficult to make decisions about usefulness of information	5.0 (2.24)	5.78 (2.39)

Note: A scale of 0 (not difficult) – 10 (very difficult).

\* Indicates  $p < 0.05$ .

\*\* Indicates  $p < 0.01$ .

**Table 5**  
Perceived mental effort from Experiment 2.

Questions	Research Task	
	Web search mean (SD)	Lib. search mean (SD)
Effort invested into searching	4.56 (1.59)	6.22 (1.30)
Thoughts put into searching	5.67 (1.22)	6.33 (1.41)
Concentration during the search**	6.44 (1.24)	8.0 (1.12)
Usefulness of search results	6.67 (2.24)	5.56 (2.30)
Importance of search	5.63 (2.14)	5.75 (2.43)

Note: A scale of 0 (no effort, no thinking, not concentrated) – 10 (a great deal of effort, a great deal of thinking, very concentrated).

\*\* Indicates  $p < 0.01$ .

( $t(8) = 3.83$ ,  $p < 0.01$ ), and the Single-Task condition and the Library Search Dual-Task condition ( $t(8) = 1.96$ ,  $p < 0.05$ ) were found to be statistically significant using the paired  $t$ -test. No significant difference was found in the reaction times for the Library Search Dual-Task condition and the Web Search Dual-Task condition using the paired  $t$ -test.

It was found that more misses and false alarms occurred during library system searching than during Web searching. The number of misses for the Library Search Dual-Task condition ( $M = 2.78$  times,  $SD = 2.54$ ) was higher than for the Web Search Dual-Task condition ( $M = 1.78$  times,  $SD = 2.33$ ). The number of false alarms was slightly higher for the Library Search Dual-Task condition ( $M = 0.44$  times,  $SD = 0.24$ ) than for the Web Search Dual-Task condition ( $M = 0.33$ ,  $SD = 0.17$ ). We cannot conclude that the subjects might have concentrated more in library system searching than in Web searching because the differences were not statistically significant.

In addition to the comparison of frequencies in misses and false alarms, we examined in what state of the searching process these errors (misses and false alarms) occurred. A particularly interesting finding is that for the Web searching condition, subjects did not miss any visual stimuli or responded to false alarms when they were scanning search results. In this condition, the majority of errors occurred when subjects were reading documents. In the library system searching condition, however, the errors were made when subjects were scanning the search results as well as reading the documents.

### 5.2.3. Analysis of search transaction logs

Transaction log analysis in Table 6 presents more insights into mental effort invested in searching. The documents in the library system referred to full catalog entries (as opposed to summary entries displayed in search results page), full text documents retrieved from online database systems, or library-created Web pages (such as library hours, information for students and information for faculty). The results showed that while subjects spent more time viewing search results in the library system (3 min 24 s) compared to the Web (2 min), and they spent more time reading full texts retrieved on the Web (9 min 14 s) compared to the library system (5 min 27 s). It was noted that more documents were read by the subjects on the Web (13.56 documents) than in the library system (12.44 documents), which was in accord with more time spent reading documents on the Web (9 min 14 s) than in the library system (5 min 27 s). They also clicked on more links (Web  $M = 28.78$ , Lib  $M = 47.89$ ), reformulated queries more often (Web  $M = 4.78$ , Lib  $M = 6.89$ ), and used more advanced features during sessions (Web  $M = 1.67$ , Lib  $M = 2.44$ ) in the library system than on the Web. See Table 6 for more results of transaction log analysis between the two systems.

### 5.2.4. Analysis of interviews

Content analysis of the exit interviews revealed primarily subjects' perceptions of IR systems, and they were consistent with subjects' responses on questionnaires. Analysis of exit interviews identified the following themes, and they were consistent with the analysis of data from Experiment 1: (1) perceptions of IR systems, (2) self-efficacy; and (3) mental effort.

**Table 6**  
Search transaction log analysis from Experiment 2.

Transaction logs	Web search Mean	Library search Mean
Time taken to complete search session (min/s)	14:21	14:26
Time taken to view search results (min/s)	2:00	3:24
Time taken to read documents (min/s)	9:14	5:27
Number of clicks per search session	28.78	47.89
Number of search results selected per session	8.67	8.67
Number of words entered in the first query	2.56	2.44
Number of words entered in average per query	3.14	2.38
Number of query reformulations during session	4.78	6.89
Number of advanced features used during session	1.67	2.44
Number of documents viewed per session	13.56	12.44

**5.2.4.1. Perceptions of IR systems.** Subjects perceived the Web as being easy and convenient, while the library system was complicated and difficult to use. Subjects pointed out different reasons for successful and unsuccessful searches in the library system searching and Web searching. Success in Web searching was uniformly attributed to using the right keywords for search queries. Similarly, failure in Web searching was attributed to using the wrong keywords. Success in library system searching was attributed to not only using the right keywords but also knowing what kind of documents were searchable in each library database (S21). S23 mentioned the need for instruction to learn how to search the system, as opposed to the “trial and error” method used to learn Web searching: “Mirlyn and the U-M Library site are very useful but you have to have a certain amount of instruction to really work them successfully.”

The results of the perceptions of IR systems were consistent with those of Experiment 1 in which subjects blamed themselves if they could not find results on the Web, and blamed the system when they could not find results in the library system. Here are some examples of how the subjects attributed their unsuccessful library searches to the complexity and difficulty of the library system.

- The Library site's terrible. It's not easy to navigate. (S16)
- On the [library] home page, there are a lot of different links to different sources...they really didn't tell you where you're going...I didn't even really know where to start. (S18)
- The Mirlyn system is kind of complicated. It's difficult to tell what you are looking at, going just by title. (S24)

Subjects also expressed differential expectations regarding the type of materials to be found in the two systems, relating them to the credibility of these materials. Libraries were perceived as containing scholarly research, “library searches give you reputable journals and stuff a lot of the time” (S19) and “the library, you're always going to know that it's going to be reliable, they wouldn't actually have books in there that were false” (S22)—while the Web provided materials that were “a little more pop culture, a little bit lighter fare” (S23). One subject compared the credibility of library materials and Web content in terms of bias:

The Library thing is more trustworthy in a way, because a lot of the stuff is not biased. On the Web, you get a lot of these blogs and online journals of people, like, I saw an anarchist's journal, and there was really right-wing and left-wing people, and they have their own agendas, so you have to take this with a grain of salt on the Web. (S19)

**5.2.4.2. Self-efficacy.** Subjects expressed greater self-confidence in using the Web compared to the library system. This confidence was clearly related to familiarity. S19, for instance, said that “I'm more comfortable using Google and stuff, just because I use it more often.” In contrast, when using the library system a common theme was that they did not know how to use it. S18 described it as “I guess it [the library system] has a lot of resources, probably, but I don't know how to use a lot of them.” Despite their familiarity with searching on the Web, subjects could not recall receiving explicit instructions on how to use Web search engines for their searches. Subjects were able to recall instances of library instruction, whether supplied by a librarian or a faculty member.

**5.2.4.3. Mental effort.** Subjects in general did not mention directly the amount of mental effort in their searches. S18, however, explicitly compared mental effort in Web and library searching, “I was concentrating a lot more and I was more stressed on the library Web site than with Google just because I didn't really know what I was doing.” One aspect of mental effort identified by subjects was that of understanding the material presented, as expressed by S17, “It [the library system] was harder to understand, harder to find the right keywords.” In addition, S24 described that “In the library it was kind of hard to figure out where to go look.”

## 6. Discussion

Although Experiments 1 and 2 employed different research methods, the results were generally consistent across both. This section highlights important findings of the study with respect to each of four research questions.

*Research Question 1: To what extent does people's investment of mental effort in online searching differ when using a Web search engine and in an online library system?*

The results of amount of invested mental effort (AIME) seemed to be mixed depending on the measurement and data collection methods. The transaction log analysis across the two experiments consistently revealed that subjects invested greater effort in library system searching than Web searching—they clicked more links, reformulated queries more often, and used advanced features more often in library system searching. The results of the Experiment 2's dual-task performance showed subjects concentrated more and made more errors by missing the visual stimuli and responding to false alarms when using the library system than they did when using the Web. Self-reports from the post-search questionnaires, however, did not present consistent findings across the two experiments. In Experiment 1, subjects reported investing more effort, putting more thoughts, and concentrating more when using the library system only when they engaged in the Product Task, but

invested about same level of mental effort when they engaged in the Research Task. Subjects participating in Experiment 2 who performed Research Task searches said that they invested more effort in searching and concentrated more during library system searches than Web searches.

Putting these mixed results together, it is fair to say that subjects invested more mental effort in library system searching than Web searching. Subjects, however, may not always be aware of their differential investment of mental effort in the two systems as shown in their responses in the self-reported post-search questionnaires. We speculate that the reasons for inconsistent responses across the two experiments are related to the Perceived Demand Characteristics (PDC) of search tasks. When presented with tasks and systems with differing PDC, task PDC is more salient than system PDC. When Experiment 1 subjects perceived the Research Task to be more demanding and difficult than the Product Task, they might have focused on the completion of the search task and not as much on system differences. Differences in system PDC became evident in the absence of task PDC, as in Experiment 2. Overall, these results indicate that the self-report method on its own might not be reliable enough to be used to measure the amount of invested mental effort (AIME) in online searching.

*Research Question 2: How do people's perceptions of these two systems and their level of self-efficacy affect their mental effort?*

In general, results from experiments 1 and 2 verified Salomon's (1983, 1984) AIME model. Subjects in our study perceived the Web as being easy and the U-M Library system as being difficult. They seemed to have strong preconceptions about the academic library system searching in which library resources were of limited scope, for example, for "homework, papers, and projects" (S12) or "scientific journals" (S04). They perceived Web searching was easy because the Web contained all kinds of information and they could "just put the keywords" (S08) and "narrow it down from broad search topics a little bit more easily" (S06). Since the early 2000s, Web search engines have been the dominant information retrieval systems (Jones, 2002). As a result, the ease and simplicity of Web searching negatively influenced library system searching as subjects in our study applied their expectations and strategies from Web searching to library system searching. Struck by the library system's relative complexity, subjects were easily confused.

Study results indicate that subjects were more confident with Web searching than library system searching because of their familiarity and experience with the Web; however, their self-confidence with Web searching did not lead them to invest more mental effort in it. According to Clark's (2001) discussion of the "inverted U" relationship, mental effort slows then stops at either exceptionally low or high self efficacy levels. Subjects in our study seemed to be "overconfident" in their Web searching ability and did not put much effort into it. For example, they entered keywords quickly, rarely reformulated search queries, viewed search results quickly, and hardly used advanced features. When searching on the Web, subjects may be comparable to drivers driving a familiar route that has become an automatic activity they do without much conscious deliberation.

*Research Question 3: To what extent does differential investment of mental effort in the two systems influence user search experience?*

Study results demonstrate that user search experience in the two systems differs in various ways, and this experience is related to their investment of mental effort. These results seem to be consistent with Salomon's (1984) research findings in which he found that children identified different attributes for failure and success of learning for television and print materials. Children attributed their failure to comprehend television programs to their own "dumbness." When asked the same about print, they targeted the difficulty of the print materials. In this online searching study, subjects expressed viewpoints about their search success or failure that differed for the two systems. Subjects were confident about their ability to search the Web, and thus, did not invest much effort into Web searching. When they could not find Web-based information, they blamed themselves, not the search engine.

When subjects used the library system for finding information, their search experience turned out to be quite different from the Web. They perceived the library to be difficult, and were more active searching the library system, clicking more, entering more words, reformulating more queries, and using more advanced features. The problem is that the subjects did not recognize these behaviors as part of the mental effort needed to improve their search results. Instead, subjects expressed their search experience in the library system using terms such as "complicated," "overwhelming," "confusing," and "hard to understand."

*Research Question 4: Which data collection method (self reports vs. dual-task performance) provides a more reliable data set for measuring mental effort in online searching?*

Among the three main categories of methods for assessing the construct of mental effort—opinion measures, dual-task techniques, and physiological measures—this study enlisted a self-reported questionnaire and a dual-task method. In opinion measures, it is presumed that the investment of effort is a voluntary process which is under the control of the learner and is available for introspection. The results of this study demonstrate that users can indeed self-report on the extent of effort that they have invested during their searches. Our findings, however, show that users might have difficulty recollecting the exact amount of effort they put into searching. Self-reports were not accurate enough to make a comparison of mental effort between two systems or two tasks. We would argue that self-reports can be reliable when they are used with other

measurements of AIME in online searching rather than being used on their own. The dual-task method is an appropriate method for measuring mental effort in online searching. We did not encounter any significant problems in implementing a second task and in having subjects engage in two tasks—searching on the IR system and responding to a visual observation task (Kim & Rieh, 2005). The dual-task method was especially useful when used along with other methods such as search transaction logs and interviews.

In addition to the usual methods used by AIME researchers, our study used search transaction logs generated by screen capturing software Camtasia. The transaction logs show users' interaction patterns with the system at every step from query formulation to viewing full text. Logging is an unobtrusive method for capturing user activities and for understanding the extent of effort subjects put into online searching. However, using transaction logs solely may provide limited empirical findings (Rieh & Xie, 2006) because researchers are unable to understand why people perform searches in certain ways and to comprehend their extent of mental effort.

## 7. Conclusion

The primary contribution of this research is the introduction of AIME (amount of invested mental effort), a concept developed in educational psychologist Salomon (1981, 1983, 1984), to the research area of online searching. AIME provides a useful framework that encompasses the essential idea of mindfulness in online searching. Even though it is widely known that people's Web searches are very simple in view of their short queries, brief sessions, rare use of advanced features, previous studies have not provided sufficient theories or empirical findings to explain why people's searches are so simple. Our study results indicate that people's simple searches are related to their a priori perceptions of search systems and their level of self-confidence toward online searching in a particular system. We are now able to speculate on the occasions when users have low AIME. When AIME is low, people put little conscious effort into searching, and thus, they are less likely to learn from their searches. Our study demonstrates that subjects' perception of systems and self-efficacy not only influence AIME in online searching but also shape the ways they experience their searches in information retrieval systems such as Web search engines and library systems.

The methodological approach taken in this research has some implications for future research in online searching behavior. While self-reported questionnaires or the dual-task technique may not be a perfect method on their own, using multiple data collection methods—questionnaires, dual-task technique, think-aloud, interviews, and transaction logs—strengthen study results. Further, multiple data methods make it possible to capture subjects' interactions with an IR system and search experience in detail. The dual-task method in particular seems to be potentially useful for assessing various aspects of online searching behavior. For instance, Gwizdka has used the dual-task method to examine mental effort expended across different stages of the online search process (e.g., Gwizdka, 2008, 2010). More research is needed on the application of the dual-task method in interactive IR studies to identify secondary tasks particularly suited for use in interactive IR experiments. Except for Dennis et al. study (2002) using an auditory secondary task, subsequent studies have employed visual secondary tasks (e.g., Gwizdka, 2008; Schmutz, Heinz, Métrailler, & Opwis, 2009). As online searching now occurs on a variety of handheld devices, such as tablets and smart phones, haptic secondary tasks may also need to be developed.

This study's suggestions for future research agenda are related to its limitations. This study focused on three primary factors—Perceived Demand Characteristics (PDC) of systems, PDC of tasks, and Perceived Self-Efficacy (PSE)—which influence the amount of invested mental effort (AIME) (Fig. 1). Future researchers can study the factors used in this study and add new ones to test their relationship to mental effort. Examples of such factors are time constraints in online searching, interface design of IR systems, and familiarity of search tasks.

Since this experiment was conducted, the U-M Library completely revamped the library's homepage, adding federated access to the library's collections. Now people can search the book collection, abstracting and indexing databases, and electronic journals and newspapers by entering their queries into a single dialog box in an interface resembling the simple interfaces of Web search engines. Reducing the rigidity and complexity of online library systems has been the impetus for the development and implementation of federated-access library systems (Antelman, Lynema, & Pace, 2006; Mischo, Bishoff, Schlembach, & German, 2012; Yang & Hofmann, 2011). A replication of this experiment that compares Web searching and federated-access library systems is recommended to determine whether the latter have leveled the playing field, that is, whether federated library systems are now comparable to web search engines in terms of the three factors this paper probes. Generally, knowing more about AIME has the potential to hasten the development of more efficient and effective IR systems that can help people find the information they need.

## References

- Allan, J., Carterette, B., & Lewis, J. (2005). When will information retrieval be "Good Enough"? In *Proceedings of the 28th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 433–440).
- Antelman, K., Lynema, E., & Pace, A. K. (2006). Toward a twenty-first century catalog. *Information Technology & Libraries*, 25(3), 128–139.
- Bandura, A. (1982). Self efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 85, 139–152.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.
- Belkin, N. J., & Muresan, G. (2004). Measuring web search effectiveness: Rutgers at interactive TREC. *Measuring web search effectiveness: The user perspectives*. Workshop conducted at the meeting of WWW 2004, New York. <<http://web.cs.dal.ca/~watters/www2004WorkShop/muresan-belkin.pdf>>.
- Cennamo, K. S. (1992–1993). Students' perceptions of the ease of learning from computers and interactive video: An exploratory study. *Journal of Educational Technology Systems*, 21(3), 251–263.

- Cennamo, K. S. (1993). Learning from video: Factors influencing learners' preconceptions and invested mental effort. *Educational Technology Research and Development*, 41(3), 33–45.
- Clark, R. E. (2001). Cognitive and motivational research issues for multimedia instruction. In R. E. Clark (Ed.), *Learning from media: Arguments, analysis, and evidence* (pp. 263–298). Greenwich, CT: Information Age Publishing.
- 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.
- Fallows, D. (2005). *Search Engine Users: Internet searchers are confident, satisfied and trusting – but they are also unaware and naïve*. Washington, DC: Pew Internet & American Life Project. <[http://www.pewinternet.org/~media/Files/Reports/2005/PIP\\_Searchengine\\_users.pdf.pdf](http://www.pewinternet.org/~media/Files/Reports/2005/PIP_Searchengine_users.pdf.pdf)>.
- Fast, K. V., & Campbell, D. G. (2004). "I still like Google": University student perceptions of searching OPACs and the Web. In *Proceedings of the 67th annual meeting of the American society for information science and technology* (Vol. 41, pp. 138–146).
- Flad, J. A. (2002). *The effects of increasing cognitive load on self-report and dual-task measures of mental effort during problem solving*. Unpublished doctoral dissertation. University of Southern California.
- Gerwe, P., & Viles, C. L. (2000). User effort in query construction and interface selection. In *Proceedings of the fifth ACM conference on digital libraries* (pp. 246–247).
- Gevens, A., Smith, M. E., Heong, H., McEvoy, L., Whitfield, S., Du, R., et al (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors*, 40(1), 79–91.
- Gimino, A. E. (2000). *Factors that influence students' investment of mental effort in academic tasks: A validation and exploratory study*. Unpublished doctoral dissertation, University of Southern California.
- Gross, M. (1995). The imposed query. *Reference Quarterly*, 35, 236–243.
- Gwizdzka, J. (2008). Revisiting search task difficulty: Behavioral and individual difference measures. In *Proceedings of the 71st annual meeting of the American society for information science and technology* (Vol. 45).
- Gwizdzka, J. (2010). Distribution of cognitive load in Web search. *Journal of the American Society for Information Science and Technology*, 61(11), 2167–2187.
- Iani, C., Gopher, D., & Lavie, P. (2004). Effects of task difficulty and invested mental effort on peripheral vasoconstriction. *Psychophysiology*, 41, 789–798.
- Jansen, J. (2010). *Online product research*. Washington, DC: Pew Internet and American Life Project. <<http://www.pewinternet.org/~media/Files/Reports/2010/PIP%20Online%20Product%20Research%20final.pdf>>.
- Jones, S. (2002). *The Internet goes to college: How students are living in the future with today's technology*. Washington, DC: Pew Internet and American Life Project. <[http://www.pewinternet.org/~media/Files/Reports/2002/PIP\\_College\\_Report.pdf.pdf](http://www.pewinternet.org/~media/Files/Reports/2002/PIP_College_Report.pdf.pdf)>.
- Kim, Y. -M., & Rieh, S. Y. (2005). Dual-task performance as a measure for mental effort in library searching and Web searching. In *Proceedings of the 68th annual meeting of the American society for information science and technology* (Vol. 42).
- Krendl, K. A. (1986). Media influence on learning: Examining the role of preconceptions. *Educational Communications and Technology Journal*, 34(4), 223–234.
- Kuhlthau, C. C. (2004). *Seeking meaning: A process approach to library and information services* (2nd ed.). Westport, CT: Libraries Unlimited.
- Kunkel, D., & Kovaric, P. (1983). Mental effort and learning from TV: Comparing expectations for PBS and commercial network programming. *Paper presented at the annual meeting of the international communication association*. ED235539.
- Markey, K. (2007). Twenty-five years of end-user searching, Part 1: Research findings. *Journal of the American Society for Information Science and Technology*, 58, 1071–1081.
- Mellon, C. (1986). Library anxiety: A grounded theory and its development. *College and Research Libraries*, 47(2), 160–165.
- Mischo, W. H., Bishoff, J., Schlembach, M. C., & German, E. M. (2012). User search activities within an academic library gateway: Implications for Webscale discovery systems. In M. Popp & D. Dallis (Eds.), *Planning and implementing resource discovery tools in academic libraries* (pp. 173–193). Hershey, PA: IGI Global.
- 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.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429–434.
- Rieh, S. Y. (2007). A Comparison of Web searching and library system searching: Perceived difficulty, self-efficacy, and effort. *Journal of the Korean Society for Information Management*, 24(2), 29–44.
- Rieh, S. Y., Hilligoss, B., & Yang, J. (2007). Toward an integrated framework of information and communication behavior: College students' information resources and media selection. In *Proceedings of the 70th annual meeting of the American society for information science and technology* (Vol. 44).
- Rieh, S. Y., & Xie, H. (2006). Analysis of multiple query reformulations on the Web: The interactive information retrieval context. *Information Processing & Management*, 42, 751–768.
- Salomon, G. (1983). The differential investment of mental effort in learning from different sources. *Educational Psychology*, 18(1), 42–50.
- 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.
- Salomon, G. (1981). Introducing AIME: The assessment of children's mental involvement with television. In K. Kelly & H. Gardner (Eds.), *New directions for child development: Viewing children through television* (pp. 89–102). San Francisco: Jossey-Bass.
- Salomon, G., & Leigh, T. (1984). Predispositions about learning from print and television. *Journal of Communication*, 34(2), 119–135.
- Schmutz, P., Heinz, S., Métrailler, Y., & Opwis, K. (2009). Cognitive load in eCommerce applications: Measurement and effects on user satisfaction. *Advances in Human-Computer Interaction*. Article ID 121494 <<http://downloads.hindawi.com/journals/ahci/2009/121494.pdf>>.
- Smith, C. L., & Kantor, P. B. (2008). User adaptation: Good results from poor systems. In *Proceedings of the 30th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 147–154).
- Spink, A., & Jansen, B. J. (2004). *Web search: Public searching of the Web*. Boston: Kluwer Academic Publishers.
- Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction*, 8(4), 351–362.
- Turpin, A. H., & Hersh, W. (2001). Why batch and user evaluations do not give the same results. In *Proceedings of the 24th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 225–231).
- Turpin, A. H., & Scholer, F. (2006). User performance versus prevision measures for simple search tasks. In *Proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 11–18).
- Veltman, J. A., & Gaillard, A. W. K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, 41(5), 656–669.
- Yang, S. Q., & Hofmann, M. A. (2011). Current generation or next generation? A study of the OPACs of the 260 academic libraries in the USA and Canada. *Library Hi Tech*, 29(2), 266–300.