Gwizdka, J., Spence, I. (2006). What Can Searching Behavior Tell Us About the Difficulty of Information Tasks? A Study of Web Navigation. Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology (ASIS&T), vol. 43. Austin, Texas, USA.

What Can Searching Behavior Tell Us About the Difficulty of Information Tasks? A Study of Web Navigation

Jacek Gwizdka

Department of Library and Information Science, Rutgers University
4 Huntington St, New Brunswick, NJ 08901, USA. assist06@gwizdka.com

Ian Spence

Department of Psychology, University of Toronto
100 St George St, Toronto, ON M5S 3G3, Canada. spence@psych.utoronto.ca

Task has been recognized as an influential factor in information seeking behavior. An increasing number of studies are concentrating on the specific characteristics of the task as independent variables to explain associated information-seeking activities. This paper examines the relationships between operational measures of information search behavior, subjectively perceived post-task difficulty and objective task complexity in the context of factual information-seeking tasks on the web. A question-driven, web-based information-finding study was conducted in a controlled experimental setting. The study participants performed nine search tasks of varying complexity. Subjective task difficulty was found to be correlated with many measures that characterize the searcher's activities. Four of those measures, the number of the unique web pages visited, the time spent on each page, the degree of deviation from the optimal path and the degree of the navigation path's linearity, were found to be good predictors of subjective task difficulty. Objective task complexity was found to affect the relative importance of those predictors and to affect subjective assessment of task difficulty.

Introduction

Information systems are designed to support users in accomplishing their tasks. Accordingly, the user's task and its characteristics have long been recognized an important factor in information seeking. In recent years, task became a focal point of many information studies. An increasing number of studies are concentrating on a specific characteristic of the task as an independent variable to explain associated information-seeking activities. A significant group among these studies has addressed task complexity or difficulty. However, the relationships between the searcher's assessed task difficulty, the objective task complexity, and the searcher's information-seeking behavior have not been studied extensively. This was the purpose of our study presented in this paper.

Related Work

The notion of task is fundamental in studies of human performance. Tasks, their characteristics, and their effects have been a traditional topic of human factors research (Wickens, et. al., 2003). In Information Science, the nature of the task has been recognized in early information seeking models (Belkin, 1980) as an influential factor in information seeking behavior. Understanding information tasks and their characteristics is widely believed to be crucial for improving the effectiveness of IR systems (Ingwersen, 1992). The increased interest in tasks is reflected in a growing body of research on information seeking.

Many studies have focused on achieving a better understanding of tasks and their characteristics (Byström, 2002; Algon, 1997; Byström & Jarvelin 1995, Kuhlthau 1992). Among more recent work are ambitious projects aspiring to address the role and effects of task in information studies in a comprehensive and systematic manner (Byström & Hansen, 2002; Kim & Soergel, 2005). Information task characteristics, and their relation to work tasks, form part of an ongoing discourse in the information science community (Byström et. al., 2004).

The effect of information-seeking task characteristics on the search process has been studied by several researchers. The type of task was used as an independent variable and the effects on information searching strategies were examined (Hsieh-Yee, 1998). Lazonder et al. (2000) examined the effect of different types of task on the search performance of novices and experts. Interaction between the type of task and user variables (e.g. cognitive style) was studied by Kim & Allen (2002). They examined the effect of different task types (a known-item search and a subject search) on searchers' activities (e.g. time taken to complete tasks, number of web documents viewed, number of search iterations, etc.) and search effectiveness. Dependence on task characteristics was discovered in most of these studies.

Task complexity has been recognized to be one of the important factors that affect information-seeking strategies (Byström, 2002; Vakkari, 1999). Complexity can been conceptualized from many different perspectives. In his interdisciplinary review of task complexity, Campbell (1988) described three general approaches to complexity: 1) psychological (subjective), 2) person-task interaction, and 3) objective (defined by task characteristics). He advocated an objective understanding of task complexity and proposed four task aspects as the factors contributing to task complexity: i) multiple possible paths to the outcome, ii) multiple outcomes, iii) conflicting interdependence among paths, iv) uncertainty linkages between paths and outcomes.

In the field of information studies, the relationship between task complexity (difficulty) and information seeking has been investigated by several researchers (Kim, 2006; Kim & Rieh, 2005; Bell & Ruthven, 2004; Byström, 2002; Vakkari, 1999; Byström & Järvelin, 1995; Saracevic & Kantor, 1988). The prevailing perspective on complexity was as a person-task interaction. The particular conceptualizations, however, differed.

Saracevic & Kantor (1988) used the number of terms in a search query as a proxy of the information task complexity. The underlying assumption was that a higher number of query terms reflects the searcher's assessment of a higher complexity of his/her information need. The higher number of query terms might also mean that it would be more difficult for the searcher's to judge relevance of documents returned by the search system.

Byström (Byström, 2002; Byström & Järvelin, 1995) construed task complexity as a prior determinability of a task by considering the degrees to which information inputs, information search process, and output of a task can be known in advance. She suggested that, in the context of information seeking, the objective view on task complexity is not necessarily relevant and rather problematic, since the objective perspective on complexity leaves the person and the context out of the task performance. Later, however, she admitted that the objective complexity is likely to have an effect on subjective assessment of task complexity.

Bell & Ruthven (2004) largely shared Byström's perspective on task complexity. In their study of web searching, task complexity was understood as the uncertainty within the information task. They suggested that task difficulty can be affected by three factors: 1) the difficulty of understanding what information is being required, 2) the difficulty of searching, 3) the difficulty of interpreting relevance.

Kim, in her study of how task influences information seeking behavior on the web (Kim, 2006), argued that task difficulty depends on individual perception, interpretation and judgment of the objective task complexity. She discussed the following contributing factors to the perceived task difficulty: the searcher's characteristics (search experience, topic knowledge), intrinsic task characteristics (specificity of target information, information sources), and intrinsic search process characteristics (assessing or navigation a website, locating information on a webpage). Problems encountered during the search process were identified by the study subjects as the main reason for their a posteriori perception of tasks as difficult.

Reaching beyond the field of information studies to the goal-setting literature (e.g., Maynard & Hackel, 1997; 2000; Steele-Johnson et. al., 2000; Mangos & Steele-Johnson, 2001), task complexity has been approached in a more comprehensive way by considering both the subjective and objective complexity and their relationship. For example, Maynard & Hackel (1997) showed that subjective and objective complexities were only modestly related and that they contributed to task performance in different ways. The authors also demonstrated that subjective task complexity was determined by objective complexity and cognitive abilities of the searcher.

The above brief overview of related literature shows a tendency in information studies towards subjective conceptualization of task complexity. Work in other research fields tends to put more emphasis on the objective view of complexity or on integration of subjective and objective perspectives. As suggested by Vakkari (2003), it would be useful for the information studies research community to enrich its understanding of phenomena related to task complexity by examining approaches from other fields. This paper explores an approach that is informed by work from other disciplines and that encompasses both subjective and objective factors. The paper focuses on the relationships between selected measures of the searcher's activities and the subjective post-task difficulty. It also explores the relationship between the objective task complexity and the subjective task difficulty.

Research Objectives

We clarify our notion of task as it is used in this paper. Task is used to mean an information task and not a work task in a wider sense (Byström & Hansen, 2002, Byström et al., 2004). Information science literature uses information task in at least two different senses: *concrete*, which refers to human actions involved in executing a task, and *abstract*, which refers to manifestation of an information seeker's problem (Marchionini, 1995). Our use of task refers to the concrete meaning of information task. Task is thus understood as a sequence of actions performed by the searcher in the process of looking for information to satisfy current information need. The task starts after information need is generated or given to the searcher, and stops when the desired information is found (or when the searcher fails to find the desired information and gives up).

The study presented in this paper is a part of a larger research program in which we aim to establish relationships between the searchers' behavior, the selected task attributes (subjective and objective), and task outcomes (such as success on information finding). Elsewhere (Gwizdka & Spence, 2005a; 2005b), we presented the results from studies which examined the relationships among measures that characterize the searcher's activities, the searcher's disorientation, and information seeking task success. In the study reported in this paper, our interest is in examining relations between the perceived post-task difficulty, the objective task complexity and measures of information search behavior in the context of factual information-seeking tasks on the web. The study was conducted with the following goals in mind:

1. to examine relationships between the searcher's activities and subjective post-task difficulty;
2. to examine which of the searcher's activities are more important in predicting subjective task difficulty;
3. to examine the effects of task ordering according to decreasing or increasing levels of objective task complexity;
4. to examine the relationship between the subjective task difficulty and the objective task complexity;
5. to examine if the relationship between the searcher's activities and post-task difficulty is affected by levels of objective task complexity.

Methodology

We conducted a question-driven, web-based information-finding study in a controlled experimental setting.

User Task

A factual information task was used. According to Morrison et al. (Morrison et al., 2001) these types of tasks belong to the most common type on the Web and account for 25% of web search activity. A factual task is defined as an information seeking task where the user seeks a specific piece of data (e.g., the name of a person or an organization, a product information, a numerical value; a date; an address; etc.). Each participant was asked to perform the same nine tasks on the University of Michigan website.

Participants

Twenty seven undergraduate psychology students (20 females and 7 males, average age 21 years) participated in this study. The average internet experience of participants was 7 years, and the average Internet use was more than 10 hours per week. The participants were recruited from the pool of first year psychology class students (PSY101) at the University of Toronto. Subjects received course credits for participating in this study.

Procedure

The study was conducted in a university laboratory on a personal desktop computer running the Microsoft Windows 2000 operating system. Each of the nine tasks was specified in the form of a question. Participants were asked to find a single web page containing factual information that provided the answer to the question. The information requested was intended to be representative of the possible questions that a typical student might ordinarily ask. All study participants came from University of Toronto and were not familiar with the University of Michigan website. Thus, we believe that the simulated information-seeking tasks used in this study were ecologically valid and authentic. Participants controlled the start and end of each task. Each of the nine questions (listed in appendix A) appeared on screen automatically after the previous task was finished. After reading each question, a starting web page (the same for all tasks – the University of Michigan home page) was automatically opened in a new browser window. Participants were asked to navigate to the target page in a single browser window without using a search engine. A Python script, running on the searcher's computer, was used to log URLs of all visited web pages, along with the timestamps and the screen coordinates of the link, the button, or the graphic, that was clicked. After each task was finished, experimenter checked if the information (web page) found by the participant was correct (Correct) and asked the participant to rate the task difficulty on a scale from 1 (easy) to 3 (difficult) (SubjDiff).

Measures

The logged time-stamped URLs were used in subsequent calculations of measures of the searcher's behavior. The measures included the number of web pages visited in a session (TotalN); the number of unique web pages visited (UniqueN); the number of web pages on the optimal path (OptPathLen); the number of times the searcher clicked on the browser's back button in each task (BackButton); the time spent by a searcher on each web page (time_per_page); and the total time on each task (time). A derived measure, "Revisits" (Tauscher & Greenberg, 1997) was calculated using the ratio of visited (N) and unique web page counts (UniqueN) by using the following equation: Revisits = 1 – UniqueN / TotalN

If we consider the individual web pages visited by searcher to be the nodes of a graph and the links actually followed by the searcher to be the graph edges, we can compute the graph properties, such as stratum (Botafogo et al., 1992). Stratum was used to characterize searcher's behavior on similar web navigation tasks by McEneaney (2001), Shih et al. (2004), Herder et al. (2004) and Juvina et al. (2004). Stratum measures how close a navigation path is to a linear ordering. This metric is based on the notion of the status, contra-status and prestige of the nodes in the path. Nodes that are hard to reach, but from which other nodes can be easily reached, have high status. Nodes that are easy to reach, but from which it is hard to reach other nodes, have high contra-status. A node's prestige is the difference between its status and contra-status. The stratum measure is defined as the sum of the absolute prestige of all nodes divided by the maximum possible value of prestige for a fully linear ordering. Stratum varies between zero and one. A value close to zero indicates a less linear navigation path; a value close to one indicates a more nearly linear navigation path (Botafogo et al., 1992).

We assessed the similarity between the user's path and the optimal path. The optimal path is the shortest path leading to the web page containing the required target information. The notion of an optimal path implies that such a path exists for an information-finding task. In our study, we also assumed that the optimal path was unique. A similarity measure was calculated based on a well-known dynamic programming procedure by Needleman and Wunsch (Needleman & Wunsch, 1970). The method uses a global sequence alignment algorithm with a non-zero gap cost and an arbitrary distance function. The non-zero gap cost was used to apply a penalty for deviation in web navigation. Our application of the N-W algorithm assumed that nodes were uniquely identified by webpage URLs composed of three parts: <host>,<path>,<query>. The distance between two nodes was calculated based on the similarity between their three-part URLs, where matching was done between each corresponding URL part (e.g., between <paths>). We derived our similarity measure — LCSlenMax — from the N-W algorithm. LCSlenMax is the length of the longest common subsequence between the user path and the optimal path divided by the length of the user path. If no gaps in the matching paths are permitted, LCSlenMax equals the number of visited pages that lie on the optimal path. Since the N-W algorithm has a non-zero gap cost, the common subsequence may contain pages which appear on only one of the paths, thus interpretation of the LCSlenMax measure is not as straightforward as in the case with no gaps.

We also introduced an "objectivized" measure of task complexity. This complexity measure was constructed to reflect the cognitive effort associated with seeking information by navigating a web site. The measure was operationalized by breaking it down into three objective factors corresponding to the elements of a task. For fact-finding web-navigation tasks such factors include:

F1. *Path Length*: the length of the navigation path leading to the target information. This factor could be interpreted as a multiplicative factor – the longer the navigation path, the more navigation choices and relevance judgments need to be made by the searcher.

F2. *Page Complexity*: the complexity of navigation choices on each web page (i.e., visual design, link labels).

F3. *Page Information Assessment*: for target web pages (i.e. pages that contain the desired information) it is the difficulty of relevance judgment, while for non-target web pages (i.e., when the desired information has not yet been found) it is the difficulty of assessing information scent (Pirolli, 2005).

This operationalization of task complexity bears some similarity to that suggested by Bell & Ruthven (2004) (presented earlier in the Related Work section) with similar second and third elements on both lists. The first element in Bell & Ruthven's list ("1. difficulty of understanding what information is being required") was not included by us since our simulated information task setting ensured that participants understood what they were asked to find in each task. The above task complexity factors were used as guidelines by three experts who rated tasks on a scale of 1 (easy), 2 (medium), to 3(difficult). Expert assigned a level of difficulty to each task (ObjDiff). The final difficulty ranking was achieved by consensus among three experts (a faculty member, an academic researcher, and a graduate student).

Independent variables are listed in Table 1. The dependent variable used in the analysis was the searcher's subjective assessment of post-task difficulty (SubjDiff).

Table 1. Independent variables and factors.

Group	Symbol	Short Description	Lower Bound	Upper Bound	Scale Type
Web page count metrics	OptPathLen	Optimal path length	1	n/a	ratio
	UniqueN	Number of unique pages	0	n/a	ratio
	TotalN	Number of total pages	1	n/a	ratio
Web page re-visit metrics	Revisits	Ratio of revisited pages to all	0	1	ratio
	BackButton	Back button use	0	n/a	ratio
Temporal metrics	time	time on task	0	n/a	ratio
	time_per_page	average speed of clicking	0	n/a	ratio
Web graph metric	Stratum	Graph linearity	0	1	ratio
Similarity to optimal path	LCSlenMax	Ratio of LCS to TotalN	0	1	ratio
Task complexity control	ObjDiff	Expert assessed task complexity	easy (1)	difficult (3)	ordinal
	QDEOrder	Order of questions	easy 1st (0)	difficult 1st (1)	nominal

Study Design

The study used a between subjects design. The tasks were administered in two different orders based on objective task complexity. In the first order, tasks were presented to participants ranging from easy to difficult. In the second order, the tasks were ordered from difficult to easy. The order of questions was changed for every other participant, and, thus, the order of tasks was a between-subjects factor (QDEOrder).

Results and Discussion

Searcher's Behavior and Subjective Post-Task Difficulty

We first sought to answer which measures of user behavior are related to user-assessed post-task difficulty and to objective task complexity. As shown in Table 2, subjective task difficulty (SubjDiff) was significantly correlated with all independent variables. This result is similar to the relationships found by Kim (2006). However, the search mechanisms used in her study were different from ours. In Kim's study, participants started their information seeking tasks from a generic Internet search engine. In our study, the information seeking task was restricted to navigation on a single large website. Our participants did not use search engines, they were seeking the target information by using the website's navigation mechanisms. Query reformulations (logged in

Kim's study) can be compared to retractions from apparently failed navigation paths (measured in our study by means of the revisit ratio and the use of the back button). Both indicate changes in the user's search strategy, and both reflect the amount of cognitive effort involved in strategy changes. The similarity between the results of these two studies suggests that the relationships between the operational measures (such as, time spent, number of pages visited, number of changes in search strategies) and the subjectively perceived task difficulty are of fundamental nature and do not depend on particular search mechanisms.

Table 2. Correlation (Spearman's rho) between searcher's assessed task difficulty (SubjDiff), objective task complexity (ObjDiff) and the independent variables.

	time on task (time)	**time per click**	**pages visited** (TotanlN)	**unique pages visited** (UniqueN)	**revisit ratio** (Revisits)	**back button use** (BackButton)	**navigation path linearity** (Stratum)	**similarity to optimal path** (LCSlenMax)
Subjective task difficulty (SubjDiff)	0.66	0.33	0.61	0.6	0.49	0.6	-0.58	-0.55
Objective task complexity (ObjDiff)	0.41	0.1^{ns}	0.44	0.47	0.27	0.31	-0.32	-0.25

^{ns} not significant; all other correlations significant at p < .01

Table 2 also shows that objective task complexity (ObjDiff) was significantly correlated with all but one (time per click) independent variables. The correlations are, however, somewhat weaker than in the case of SubjDiff. Subjective task difficulty (SubjDiff) and objective task complexity (ObjDiff) were correlated with each other, however only modestly (r=0.41, p<.01). This indicates that perhaps the searcher's perceptions of task complexity do not only reflect the inherent complexity of the task, but also some other variables. The other variables that likely mediate the relationship between subjective and objective task complexity are individual differences among the searchers. Such a relationship was shown, for example, by Kim (2006) and by Maynard & Hackel (1997). The latter researchers found that perceptions of task difficulty were affected by objective task complexity and by cognitive ability and motivation.

A non-parametric Kruskal-Wallis test for N-independent samples was used to examine if the independent variables differed for different level of task difficulty. All variables from Table 2 were found to differ significantly for different levels of task difficulty (p<.0001). Figures 1 to 3 illustrate these differences graphically.

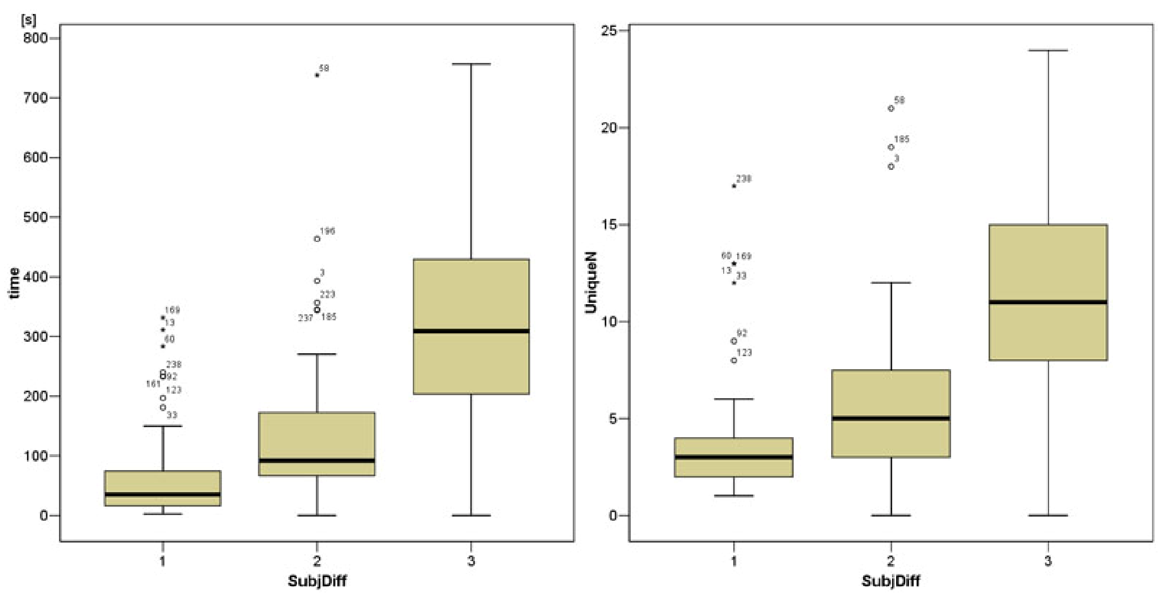

Figure 1. Time on task (time) and the number of unique pages visited by searchers (UniqueN) shown for different levels of task difficulty (SubjDiff, 1=easy, 2=medium, 3=difficult).

The more time searchers spent on the task and the more web pages they visited, the more difficult they assessed the information seeking task. This is hardly surprising since, time and the number of visited web pages both reflect effort expended on the task. The number of visited pages is related to the number clicks, and to the number of relevance judgments and navigation decisions made by searchers, thus it is an indication of physical and cognitive effort .

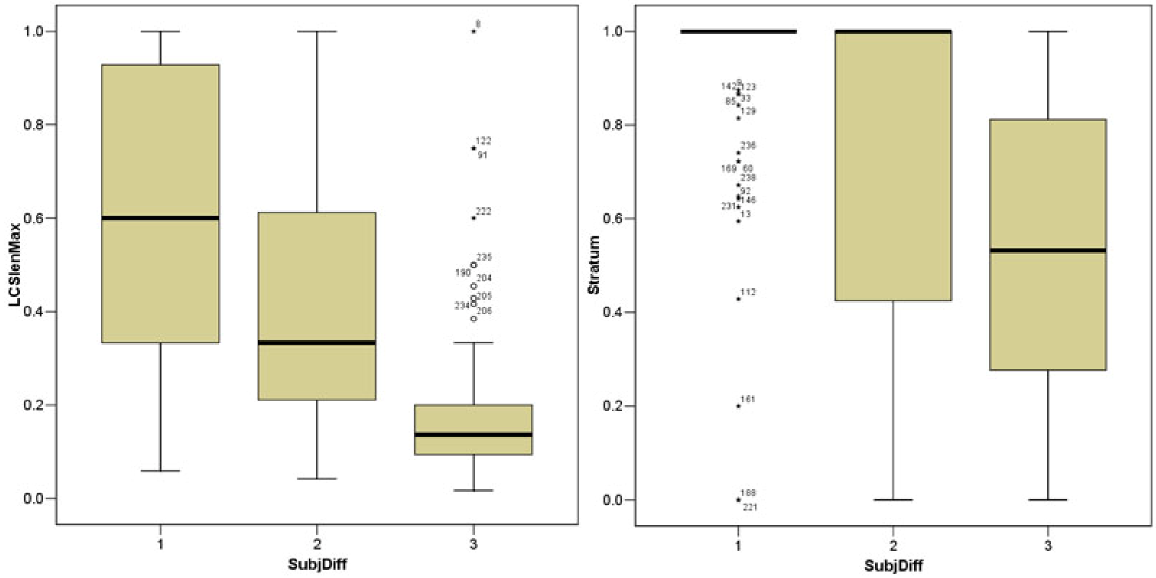

Figure 2. Similarity to the optimal path (LCSlenMax) and linearity of navigation path (Stratum) shown for different levels of task difficulty (SubjDiff, 1=easy, 2=medium, 3=difficult).

Higher similarity to the optimal path (LCSlenMax closer to 1) and a more linear navigation path (Stratum closer to 1) were associated with tasks perceived by the searchers as easy, while lower similarity to the optimal path (LCSlenMax closer to 0) and a less linear navigation path (Stratum closer to 0), were associated with tasks perceived by the searchers as difficult.

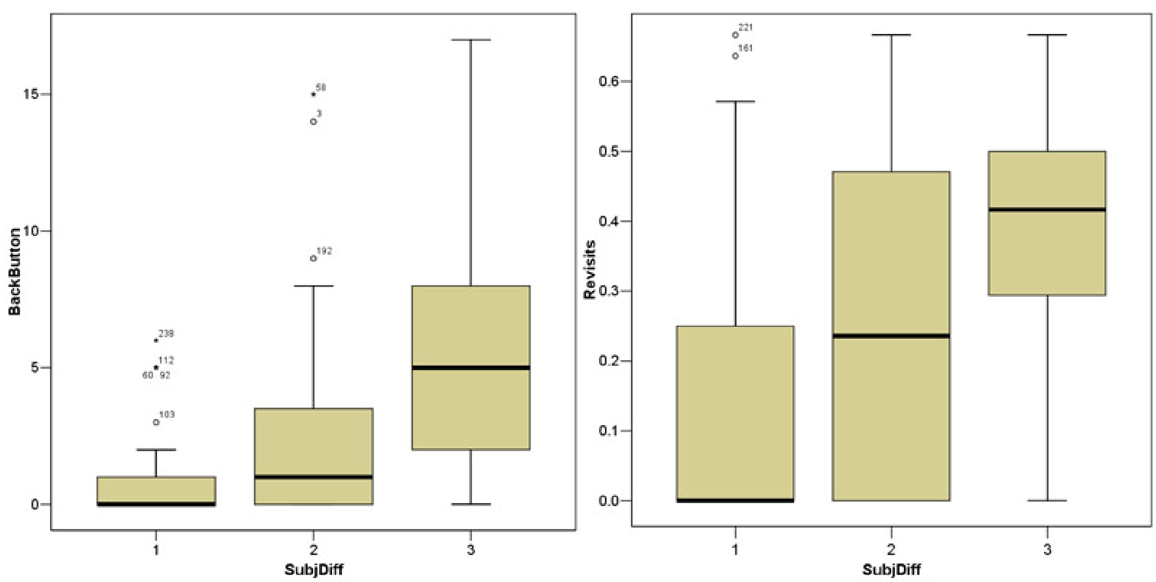

Figure 3. Back button use and the revisit ratio (Revisits) shown for different levels of task difficulty (SubjDiff, 1=easy, 2=medium, 3=difficult).

Higher back button use and larger revisit ratio correspond to situations when users retracted from their current navigation path and tried other paths. The retraction indicates a change in search strategy and associated cognitive effort. Information seeking tasks involving more strategy changes were perceived by users as more difficult.

Multiple regression was used to find a set of independent variables that best predicts subjective task difficulty, and to establish the relative importance of the independent variables. Regression models with an increasing number of independent variables were systematically examined until a model that explained the most variance with the highest number of significant predictors was found. The best single-variable model included time spent on a task as the significant predictor of the perceived task difficulty. The time variable (and the model) accounted for 36% of variance in task difficulty. The second best single-variable model included UniqueN and explained 35% of variance in task difficulty. It is worth noting that time was not present in the best higher-order regression models. The "optimal"[1] model contained four independent variables: UniqueN, time_per_click, LCSlenMax, and Stratum. The four variables accounted together for 48% of variance in the searcher perceived post-task difficulty. We obtained the following model with parameter estimates (b – coefficients):

$$\text{Predicted SubjDiff} = .05*\text{UniqueN} + .03*\text{time\_per\_click} - .51*\text{LCSlenMax} - .48*\text{Stratum} + 1.61; \quad R^2=.48$$

Table 3. Predictors of subjective task difficulty (SubjDiff) in the optimal model.

Component group	**Independent variable**	**Standardized Beta coeff.**	**Incremental contribution to variance explained**	**Unique variance explained** (based on Type II SS)
Search effort	UniqueN****	.34	+36%	6%
Navigational speed	time_per_click****	.27	+7.3%	7%
Search efficiency	LCSlenMax**	-.19	+3.3%	2%
	Stratum**	-.18	+1.9%	2%

** $p < .01$, **** $p < .00001$

Tasks tended to be a posteriori perceived as more difficult if more unique web pages were visited, more time was spent on each visited, or re-visited, web page, and the navigation path was farther away from the optimal path, as well as it was less linear. The four variables included in the model can be considered as belonging to three groups:

P1. *Search Effort* : the number of unique pages visited by the searcher in each task;

P2. *Navigational Speed*: how much time on the average the searcher spent on each page;

P3. *Search Efficiency*: deviations from the optimal path (LCSlenMax), and bushy navigation (Stratum).

General search effort was the most important component and accounted for largest part of task difficulty variance (36%). The speed component explained accounted for additional 7.3% of variance. These first two components were about equally important (as assessed by their unique contribution to variance). The two variables in the efficiency component accounted for over 5% of incremental contribution to variance. Both of them were equally important in this model (as assessed by their unique contribution to variance).

The predictor groups in this model are related to the objective task complexity factors (F1-F3, described earlier in the sub-section Measures). Search Effort (P1) corresponds to Path Length (F1), while Navigational Speed (P2) is related to both Page Complexity (F2) and Page Information Assessment (F3). Search Efficiency (P3) is related to Path Length (F1), but it also captures differences among the searchers' web navigation strategies (i.e., how individuals navigate a website). Establishing such associations may be somewhat misleading, however. We need to remember that factors F were established a priori to reflect the cognitive effort associated with seeking information on a website. These factors represent an idealized case of a well informed expert searching for information. They were established without specific searcher in mind. For example, Path Length (F1) represents length of the optimal navigation path that leads to the desired information, while Search Effort (P1) represents the number of unique web pages actually visited by an individual searching for the desired information. Thus factors F and predictors P are linked by a particular individual who performs search. The difference between factors F and predictors P represents the difference between the expected cognitive effort and the actual cognitive effort. The relationship between both groups is thus mediated by searcher's characteristics, such as experience, domain knowledge, and abilities (e.g., cognitive). Similarly to the earlier discussion of the correlational results from Table 2, this analysis suggests importance of considering individual differences among searchers in the studies of task complexity.

[1] The model was optimal in a sense of maximizing R^2 with the highest number of significant predictors.

Effects of Task Order

One of the study goals was to examine the effects of task ordering according to decreasing or increasing levels of task complexity. The intuition underlying this research objective was derived from the principle of Cumulative Advantage or “Success-Breeds-Success”, first proposed by Simon (1957). However, we did not find any effects of the two different task orders: “easy-first” and “difficult-first” (QDEOrder). Our intuition that ordering tasks by difficulty influences performance was not confirmed. However, such effect may still exist. It may be small and it may be confounded by the effects of individual differences, which were not measured in this study.

Effects Task Complexity

We examined if the importance of the variables that predict subjective task difficulty (Table 3) was mediated by the objective task complexity. The optimal regression model was calculated for the two extreme cases of task complexity: easy (ObjDiff=1) and difficult (ObjDiff=3). As shown in Table 4, the significance and the relative importance of the independent variables was different for each level of task complexity.

Table 4. Differences in the optimal regression model for different levels of task complexity.

Component Group	**Independent variable**	**Unique variance of subjective difficulty explained** (based on Type II SS)		
		All tasks* ObjDiff = all levels	**Simple tasks** ObjDiff = easy (1)	**Complex tasks** ObjDiff = difficult (3)
Search effort	UniqueN****	6%	not significant	not significant
Navigational speed	time_per_click****	7%	4%	not significant
Search efficiency	LCSlenMax**	2%	not significant	9%
	Stratum**	2%	8%	not significant

*repeated from Table 3.

Low complexity tasks were characterized by shorter optimal paths (2-3 “clicks”) and, by definition, required less cognitive effort (as assessed by experts). For these tasks, search efficiency (as expressed by linearity of the navigation path) and navigational speed remained significant predictors of subjective task difficulty. Deviations from a linear path and “slow clicking” seem to be more important factors affecting perceived task difficulty for simple tasks. We can presume that for these simple tasks the average time spent on each page reflects the cognitive effort involved in deciding which link to follow next in order to find the needed information. That is, the effort involved in assessing information scent of proximal cues represented by links (Pirolli, 2005).

High complexity tasks were characterized by longer optimal paths (5-6 “clicks”) and, by definition, required more cognitive effort (as assessed by experts). For these tasks, search efficiency (as expressed by similarity to the optimal navigation path) remained significant predictor of subjective task difficulty. Deviations from the optimal navigation path seem to be more important factor affecting subjective task difficulty for more complex tasks.

A Kruskal-Wallis test for N-independent samples was performed to check if the independent variables differed for different level of task difficulty. For low subjective task difficulty (SubjDiff=1), all variables, except the degree of similarity to the optimal path (LCSlenMax), were found to be significantly different for the two extreme levels of task complexity. The number of visited unique pages was also different for the medium level of subjective task difficulty (SubjDiff=2). Figure 4 illustrates the selected differences graphically.

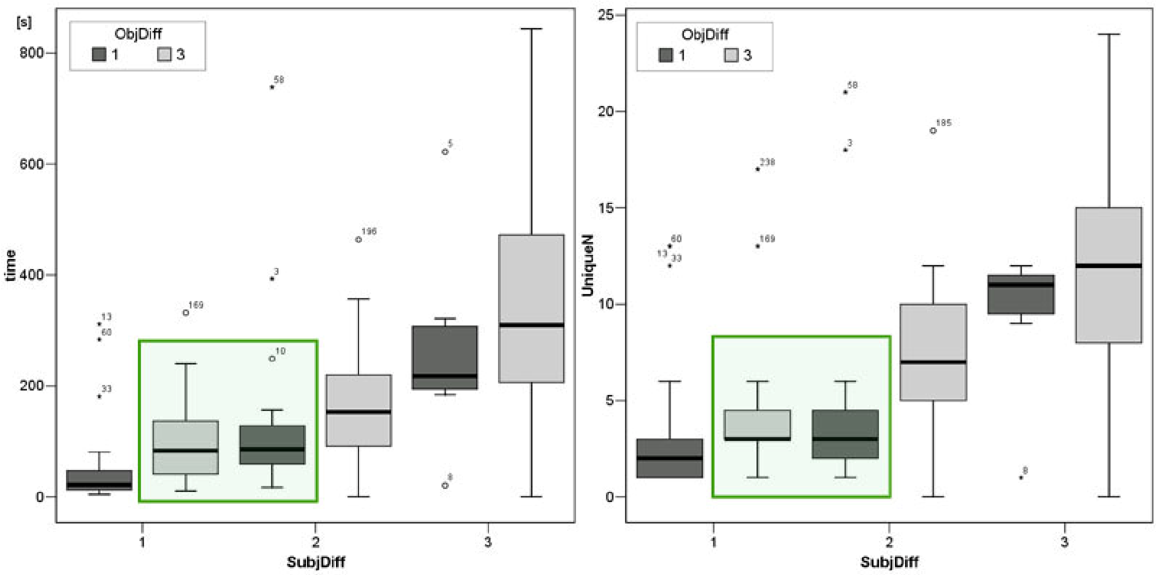

Figure 4. The differences in time on task and the number of unique pages visited by searchers (UniqueN) for the two extreme levels of task complexity (ObjDiff =1 and 3) shown for three different levels of task difficulty (SubjDiff, 1=easy, 2=medium, 3=difficult).

In two cases highlighted in Figure 4, the range of a variable (time or UniqueN) was quite similar for different combinations of subjective task difficulty and objective task complexity. That is, the range of a variable was quite similar for SubjDiff=1 & ObjDiff=3 as it was for SubjDiff=2 & ObjDiff=1. This suggests that the searcher's assessed task difficulty was not independent of objective task complexity. Complex tasks (ObjDiff=3) were perceived as low difficulty tasks (SubjDifff=1), while simple tasks (ObjDiff=1), characterized by the same range of time and UniqueN, as the complex tasks, were perceived as medium difficulty tasks. Thus, UniqueN and time could not be used to fully explain perceived task difficulty. This demonstrates the complexity of the relationships among operational measures, task characteristics (such as complexity), and perceived task characteristics (such as subjective task difficulty). One implication is that proper description of task characteristics may require a relatively large number of operational measures of the searcher's activities, or it may be limited to very specific contexts.

SUMMARY AND CONCLUSIONS

This research has focused on the relationships between operational measures of information search behavior, objective task complexity, and subjectively perceived post-task difficulty in the context of factual information-seeking tasks on the web. Higher search effort, lower navigational speed and lower search efficiency were found to be good predictors of subjective post-task difficulty. Task complexity was introduced as an "objectified" measure that could be conceptualized as a proxy of inherent task structure, when the task is performed by an expert on a given information system. Task complexity was found to affect the relative importance of the predictors of subjective task difficulty and to affect subjective judgments of task difficulty.

Comparison of our findings and those of Kim (2006) suggests that the relationships between the selected operational measures of the searcher's behavior (such as, time spent, number of pages visited, number of changes in search strategies) and the subjectively perceived task difficulty are possibly of fundamental nature and do not depend on particular search mechanisms.

Our analysis of correlational relationships indicates that the searcher's perceptions of task complexity do not only reflect the inherent complexity of the task, but also some other variables. As other researchers, we suggest that individual differences among the users mediate the relationships between objective task complexity and subjective task difficulty. That is, the searcher's perception of task difficulty will depend on such individual factors as the searcher's experience, domain knowledge, verbal ability, other cognitive abilities, and also on motivation and affect. Individuals who are low on certain abilities will perceive simpler tasks as more difficult, while those higher on the abilities will perceive more complex tasks as easier. This kind of relationship was shown by Maynard & Hackel (1997) on a scheduling task. We have no reason to believe that such relationship be different for information tasks. Ultimately, we argue that understanding the searcher's subjective assessment of task difficulty requires knowledge of the objective task complexity and the searcher's characteristics. Only then we can attempt to explain the relationships between the searchers perception of task difficulty and the searchers performance on the task. Figure 5 depicts the discussed relationships.

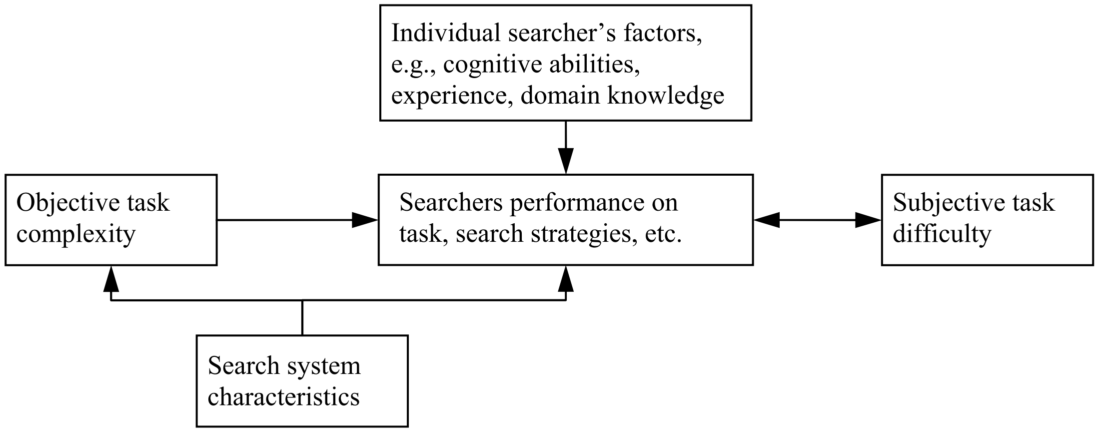

Figure 5. The emerging model of relationships between the subjective and objective task complexity.

Characterization of information tasks that are performed in different contexts and by different groups of people holds a promise for the creation of personalized information retrieval systems. For example, modeling tasks in terms of operational measures of the searcher's activities may enable us to build real-time predictive task models and thus to create IR systems that adapt to the user. To be successful, however, such systems would need to be able to assess the user's characteristics from the user's activities.

Tasks are currently at the forefront of information studies. In this research we aim to gain a better understanding of task characteristics and their relation to the objective measures of the searchers' behavior. We hope to contribute to the on-going efforts in the information studies community.

ACKNOWLEDGMENTS

We thank Neetu Manwani and Alexandra Simion for their assistance in running the study sessions and Ken Seergobin for the logging script. This research was supported by CITO (Communications and Information Technology Ontario), a division of Ontario Centres of Excellence (OCE) Inc.

REFERENCES

Algon, J. (1997). Classification of tasks, steps, and information-related behaviors of individuals on project teams. *Proceedings of International Conference on Research in Information Needs, Seeking & Use in Different Contexts*, 205-221.

Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, 5, 133-143.

Bell, D. & Ruthven, I. (2004). Searcher's Assessments of Task Complexity for Web Searching. *Advances in Information Retrieval, 26th European Conference on IR Research*, ECIR 2004 (pp. 57-71). Springer.

Botafogo, R. A., Rivlin, E., & Shneiderman, B. (1992). Structural analysis of hypertexts: Identifying hierarchies and useful metrics. *ACM Transactions on Information Systems*, 10, 142-180.

Byström, K. & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing and Management*, 31, 191-213.

Byström, K. (2002). Information and information sources in tasks of varying complexity. *Journal of the American Society for Information Science and Technology*, 53, 581-591.

Byström, K. & Hansen, P. (2002). Work tasks as unit for analysis in information seeking and retrieval studies. *Proceedings of the 4th International Conference on Conceptions of Library & Information Science* (CoLIS4), 239-252.

Byström, K. et. al. (2004). Conceptions of task as a methodological issue: Scandinavians on information seeking and retrieval research (SIG USE). *Proceedings of the American Society for Information Science and Technology*, 41, 577-579.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of Management Review*, 13, 40-52.

Gwizdka, J., Spence, I., (2005a). Predicting Outcomes of Web Navigation. *Poster presented at WWW'2005*, Chiba, Japan.

Gwizdka, J., Spence, I. (2005b). Indirect Assessment of Web Navigation Success. *Extended Abstracts of ACM Conference on Human Factors in Computing Systems CHI'2005*. ACM Press. 1427 - 1430.

Hackman, J. R. (1969). Toward understanding the role of tasks in behavioral research. *Acta psychologica*, 31, 97-128.

Hansen, P. (1999). User interface design for IR interaction: A task-oriented approach. *Proceedings of the Third International Conference on Conceptions of Library and Information Science (CoLIS 3).* 191-205.

Herder, E. (2003). Revisitation Patterns and Disorientation. *Proceedings of the German Workshop on Adaptivity and User Modeling in Interactive Systems ABIS 2003*, 291-294.

Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science*, 44, 161-174.

Ingwersen, P. (1992). *Information Retrieval Interaction*.

Juvina, I. & van Oostendorp, H. (2004). Individual Differences and Behavioral Aspects Involved in Modeling Web Navigation. *Lecture Notes In Computer Science*, 3196, 77-95.

Kim, J.-H. (2006). Task as a Predictable Indicator for Information Seeking Behavior on the Web. *An unpublished dissertation*. SCILS, Rutgers University.

Kim, Y. M. & Rieh, S. Y. (2005). Dual-Task Performance as a Measure for Mental Effort in Library Searching and Web Searching. *Proceedings of the 68th Annual Meeting of the American Society for Information Science and Technology*.

Kim, S. & Soergel, D. (2005). Selecting and Measuring Task Characteristics as Independent Variables. *Proceedings of the 68th Annual Meeting of the American Society for Information Science and Technology*.

Kuhlthau, C. C. (1993). A principle of uncertainty for information seeking. *Journal of Documentation*, 49, 339-355.

Lazonder, A. W., Biemans, H. J. A., & Wopereis, I. G. J. H. (2000). Differences between novice and experienced users in searching information on the World Wide Web. *Journal of the American Society for Information Science*, 51, 576-581.

Mangos, P. M. & Steele-Johnson, D. (2001). The Role of Subjective Task Complexity in Goal Orientation, Self-Efficacy, and Performance Relations. *Human Performance*, 14, 169-186.

Marchionini, G. (1995). *Information seeking in electronic environments*. Cambridge Series on HCI, 9.

Maynard, D.C., & Hakel, M. D.(1997). Effects of objective and subjective task complexity on performance. *Human Performance*, 10, 303-330.

McEneaney, J. E. (2001). Graphic and numerical methods to assess navigation in hypertext. *International Journal of Human Computer Studies*, 55, 761-786.

Morrison, J., Pirolli, P., & Card, S. K. (2001). A taxonomic analysis of what world wide web activities significantly impact people's decisions and actions. *Proceedings of CHI' 2001 - Extended abstracts*, 161-16).

Needleman, S. B. & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of molecular biology*, 48, 443-453.

Pirolli, P. (2005). Rational analyses of information foraging on the web, *Cognitive Science,* 29 (3), 343-373.

Saracevic, T. & Kantor, P. (1988). A study of information seeking and retrieving. III. Searchers, searches, and overlap. *Journal of the American Society for Information Science*, 39, 197-216.

Shih, P.-C., Mate, R., Sanchez, F., & Munoz, D. (2004). Quantifying user-navigation patterns: a methodology proposal. In Poster presented at the 28th International Congress of Psychology in Bejing Bejing 2004.

Simon, H. A. (1957). *Models of man: Social and rational*. New York: John Wiley & Sons.

Steele-Johnson, D., Beauregard, R. S., Hoover, P., & Schmidt, A. M. (2000). Goal orientation and task demand effects on motivation, affect, and performance. *Journal of Applied Psychology*, 85, 724-738.

Tauscher, L. & Greenberg, S. (1997). How people revisit web pages: Empirical findings and implications for the design of history systems. *International Journal of Human Computer Studies*, 47, 97-137.

Vakkari, P. (1999). Task complexity, problem structure and information actions. Integrating studies on information seeking and retrieval. *Information Processing and Management*, 35, 819-837.

Vakkari, P. (2003). Task-based information searching. *Annual Review of Information Science and Technology*, 37, 413-464.

White, R. W., Jose, J. M., & Ruthven, I. (2003). A task-oriented study on the influencing effects of query-biased summarisation in web searching. *Information Processing and Management*, 39, 707-733.

Wickens, C. D., Lee, J., Liu, Y. D., & Gordon-Becker, S. (2003). *Introduction to Human Factors Engineering* (2nd Ed.). Prentice Hall.

Appendix A.

Table 5. Questions used in the study.

No.	Question text
1	What is a Wolverine?
2	Find a listing of all recreational athletic facilities at UM
3	Find a listing of museums, galleries (including gardens etc.) at UM
4	What are the swimming pool hours?
5	What is the Rackham school?
6	Find all current job listings available to UM students
7	Find full term tuition for NON-Michigan residents in the graduate school of medicine (MD)
8	How many times has the UM ice hockey team won the NCAA championship?
9	What is the name of the Chair of the Department of Psychology?