

# The Effect of Task Domain on Search

Elaine G. Toms<sup>1</sup> Luanne Freund<sup>1</sup>

<sup>1</sup>University of Toronto  
Faculty of Information Studies  
Toronto, ON  
416-978-7802

{toms, freund, bartlett}@fis.utoronto.ca

Richard Kopak<sup>2</sup> Joan C. Bartlett<sup>1</sup>

<sup>2</sup>University of British Columbia  
SLAIS  
Vancouver, BC  
604-822-2898

rkopak@interchange.ubc.ca

## 1. Abstract

The objective of our work is to assess how search interfaces can be personalized according to domain-specific needs. In this study we investigated how people search within diverse domains (consumer health, shopping, travel and general research) to identify differences in searching needs. A mixed method research design was used to observe forty-eight participants interacting with a modified version of Google to complete four search problems from four domains. Results indicated significant differences by domain. Furthermore, analysis of verbal protocols identified specific areas to be addressed in the design of new search interfaces.

## 2. Introduction

The quest for information is generally acknowledged to be a complex task, and for users faced with the Web's unstructured and heterogeneous collection this is certainly a truism. In designing information retrieval systems for the Web, we need to deconstruct tasks, strategies, and behaviours that make up "Web searching" in order to better support users' goals [12, 16, 18].

To that end we have undertaken a user-centred study to better understand the impact of task domain on Web searching behaviour. This study was experimental but also exploratory. We examined participant behavior both quantitatively and qualitatively by task domain (consumer health, general research, shopping and travel). The domains selected for examination in this study have been identified as among the top Internet search task domains [16, 21]. Finally, we

compared the outcomes using a series of efficiency, effectiveness, and satisfaction metrics as well as verbal protocol data to gain a rich description of strategies and tactics used and rationales for observed search behaviour. Our goal was to identify the challenges and distinguishing characteristics of searching within these domains, in order to inform the design of more supportive Web search systems.

## 3. Previous Work

Studies of search tend to be of two types: a) design and development of effective algorithms for matching and ranking, and b) the study of user search behaviours. Among the latter are included generic Web searching [11], different user groups [7, 14] and task specifications [10, 17]. Little research has compared searching across task domains, although domain is considered to be an important element in models of Web searching [12, 16, 18]. This is surprising considering the importance of domain in established commercial bibliographic databases [3].

Studies of searching within particular domains suggest that differences in search behaviour do exist. Within the consumer health domain, key issues are information quality, reliability and accessibility, which vary widely across different websites and search engines [13, 23]. The source of consumer health information, while often considered an indicator of reliability, is not always an accurate predictor [15]. Searchers tend to compare the information from a Web search

against previous knowledge, or verify the findings with a health professional [19].

Use of the Internet for research is widespread, as evidenced by the large number of online Web research guides [1]. Studies point to the potential of the Web for academic research [2], but there are few user studies of Web searching behaviour in this specific context. Fidel and colleagues [5] found that high school students using the Web for research tended to look for one "landmark" page that would serve as a research base and made quick decisions regarding relevance. General studies of Web search tasks suggest that subject keyword searching is prevalent [10] and that domain and search expertise are important factors in search success [9, 14].

E-commerce is one of the fastest growing domains on the Web [20], and many consumers use general purpose search engines for shopping tasks. Spiteri [22] suggested existing search directories do not effectively support the shopping paradigm, due to terminological ambiguity, confusing categorization, and limited coverage. Spink and Gunar [20] found that shopping queries, tend to be slightly longer than general queries, and commonly include the names of companies, stores or products.

Although typically functional, travel information is often sought to satisfy experiential, creative, and aesthetic needs [1]. Distinct strategies of information seeking have also been noted within the travel domain [6]. Fodness and Murray [7] found that search strategy is affected by a number of "contingencies" including the purpose of the trip, composition of the traveling party, and socio-economic status. Gursoy [8] found travel information search strategies dependent on familiarity with destination or product and the level of expertise in search.

Following from these domain specific studies, we asked the following question: *is a generic search interface (such as that presented by Web search engines like Google) sufficient for effective searching across all domains?*

## 4. Methodology

### 4.1 Participants

The 48 participants (29 f and 19 m) were adult members of the general public (including but not limited to the university community) who had used the Web, and who may have had some training, but had not taken a professional search course. They ranged in age from 18-20 to over 65 years; 80% were under 35. Most had university level education, mainly at the bachelor (38%) or masters (30%) level, predominantly from the humanities or social sciences. About half were students; the remainder was from a diverse range of occupations. All but one participant reported searching the Web on a daily or weekly basis. Overall, the participants were a relatively young, educated group who were experienced, but moderate, Web searchers.

### 4.2 Search Interface

Google was used because of its acknowledged popularity, but the standard interface was modified to include both the search box/button, and the Google top-level category list (directory). The resulting screen retained Google's simplicity, but includes an instruction to enter either a query in the search box or select a category from the directory (See Figure 1). Beyond this initial page, the standard Google interface was retained.

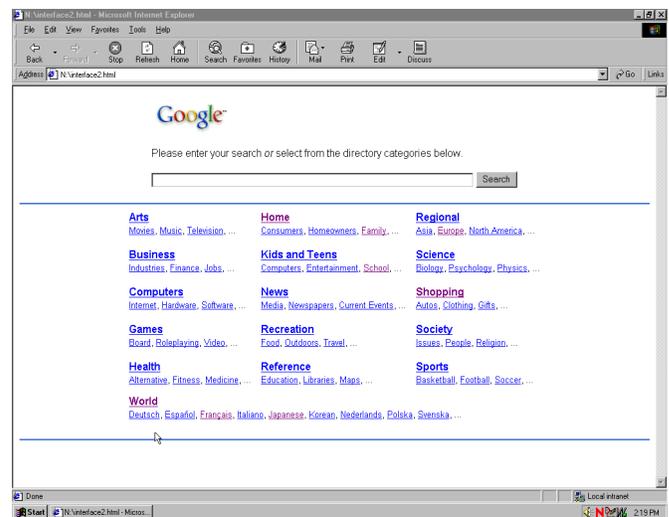


Figure 1. Modified Google Search Interface

### 4.3 Search Problems

The sixteen problem statements used in the study came from four domains: Consumer Health,

Research, Travel and Shopping. Of the 16 problems, half were fully specified and half were partially specified so that participants could personalize them. The problems statements were those used in the Interactive Track of the 10<sup>th</sup> Annual Text Retrieval Conference [<http://trec.nist.gov/>].

#### 4.4 Procedure

The participants were recruited in Toronto and Vancouver. Each was assigned four search problems, one from each of the four domains. Two of the four problems contained specific questions or imperatives that the participant was asked to respond to by finding relevant Web pages. For the remaining two problems, questions were partially specified, with a fill-in-the-blank portion allowing the participants to personalize the problem. We used a modified Latin squares method to distribute the problems among the participants.

Data was collected using four mechanisms:

1. Questionnaires for demographics, and pre- and post-search evaluations.
2. Audio-tape for all semi-structured interviews.
3. Transaction logging (*WinWhatWhere* software) to record titles and URLs of sites visited, keystrokes entered, and to time stamp data.
4. Screen capture software (Lotus Screencam) to video record the user processes.

Each session lasted approximately two hours. During this time, participants:

- completed a demographics and Web/search experience questionnaire.
- were assigned four search problems in sequence.

For each problem, the participant

- a) completed a pre-search questionnaire to gauge familiarity with the topic,
- b) searched for the topic using the Web interface, responded to a post-search questionnaire about the search process and their satisfaction with the results, and
- c) using a 'talk-after' verbal protocol, described the process and challenges met in doing the search while viewing a screen-capture video of the search. During these audio-taped retrospective interviews, we tried to elicit the decision-making process used at each stage in the search process.

- responded to a series of questions intended to elicit comments about the challenges of searching the WWW.

#### 4.5 Data Analysis

Data from questionnaires and the transaction logs were combined. The transaction logs were manually coded for search state (e.g., query use, category selection, hit list selection, view a URL), by reviewing simultaneously the screen capture files and the transaction files. The additional coding made it possible to identify the path taken in each search, to determine the amount of time spent at each state, and to identify the rank position on a hit list page for each selected URL. Quantitative data for each of the factors were analyzed using SPSS multivariate and univariate general linear model. The audio-tapes of interviews were transcribed and coded in Nvivo to tag discussions of the various search states, challenges, and problems encountered. To supplement user perception metrics, the results from each search were assessed. Each page identified as relevant by the participant was rated for topical relevance (aboutness), and the set of pages returned for each task was rated for completeness by expert evaluators.

### Results

#### 5.1 Summary

The 48 participants spent an average of seven minutes working on each problem. They used the search box for approximately 66% of the problems and selected directory categories for the remainder. On average, they examined about five URLs and about six links within each of those URLs. They tended to select the fourth item on a hitlist and, on average, examined approximately two pages of hitlists. Participants reported little familiarity with the assigned topics. On an ascending five-point scale, they rated the ease of finding the answer (3.7), and their satisfaction (3.6) with the process of finding their answer. Similarly, they assessed the degree of certainty with which they found their answer (3.9). The external assessment of the completeness of the task averaged 4.4.

#### 5.2 Search Domain

The differences among the domains were assessed using several efficiency, effectiveness and

satisfaction metrics as well as data from the verbal protocols.

### 5.2.1 Efficiency

Efficiency was measured using time taken at various search states and the number of instances of each search state. Few significant differences existed among the four domains. There were, however, differences in post hoc Bonferroni adjusted tests. Participants did more printing in Research than in Shopping ( $p=.035$ ) and spent less time in Categories while responding to a Research task than a Travel task ( $p=.010$ ). Those doing Shopping and Travel tasks spent significantly more time looking at the content of a Website than those doing Consumer Health and Research (see Figure 2).

### 5.2.2 Effectiveness

Effectiveness was assessed in three ways: a) how certain the participant was that an appropriate answer had been found; b) how complete the solution was according to an external rater; and c) the extent to which the pages selected were about the topic. Post hoc Bonferroni adjusted tests indicated differences. Pages retrieved for Travel tended to be less about the topic than those retrieved for Research ( $p=.027$ ) with a lesser effect when compared to Shopping ( $p=.074$ ). When doing the Shopping problem, participants were less certain that they found good answers ( $p=.031$ ) in comparison to other domains.

### 5.2.3 Satisfaction

After each task participants rated their perception of the task: a) ease of use, b) amount of time allotted to do the task, c) how much their familiarity with the task contributed to their success, and d) their overall rating. The type of domain had an effect on user perception. In general, participants found the Shopping task more difficult and less satisfying than the other tasks, rating ease at an average of 3.3 and satisfaction at 3.1 on a five-point scale.

### 5.2.4 Strategy Used

Search paths were analyzed according to strategy. To start, participants elected to either submit a query or select a directory category, but could switch between query or directory at any time later. Some participants, for example, used a single tactic such as query only, while some used

novel strategies that combined queries and categories.

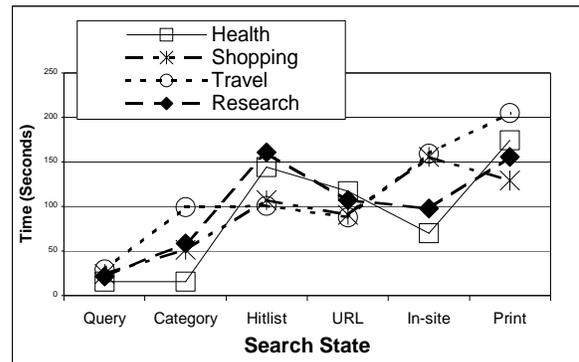


Figure 2. Time in Each Search State

As illustrated in Figure 3, those doing Consumer Health and Research searches tended to follow a standard pattern: most were completed primarily with queries and rarely with categories. Categories, however, tended to be used by those working in the Travel and Shopping domains. Because of the variation in use, and low overall use of categories, no statistical assessment concerning the differences could be performed, although a trend seems apparent.

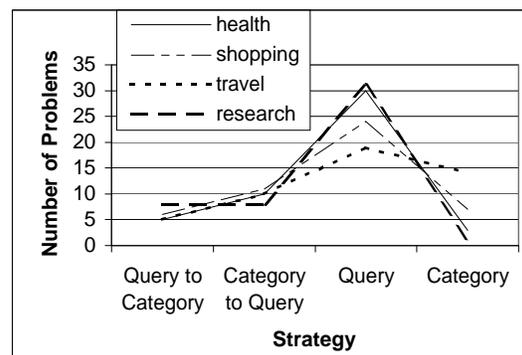


Figure 3. Strategy Used

A caveat of this result is the effect of an inherent bias in the interface towards the query. Although the start page contained both categories and a query box, once a query was used, the categories had to be sought out. On the other hand, the query box appears at the top of all subsequent category pages, and on all hitlist pages.

### 5.2.5 Search Processes

Following each search, the screen capture was replayed for the participant and a semi-structured

interview ensued. The interview used a series of questions to isolate decision-making processes and strategies at each decision point in the search. Data for 32 participants have been analyzed, and the 128 interviews (4 search problems X 32 participants) were assessed by search state (see the x axis in Figure 2) and strategy. Verbal protocols were grouped and analyzed first by information problem and the results were integrated by task domain.

#### **5.2.5.1 Consumer Health**

Consumer health searches were typically goal focused and selective. The key factors affecting selection of information sources were: trustworthiness, orientation (professional vs. consumer), and depth (lists vs. narratives). Participants reported being more rigorous in this domain than in others: “[for a health question, I] would have kept going and found something more reputable”(P32). Participants often attempted to identify the source of a website from the hitlist in order to select more trustworthy sources. University, government and association sites were generally considered to be authoritative; whereas, commercial sites were considered to be biased. The target audience of the website was more difficult to gauge from the hitlist, and participants often resorted to trying to identify this from the terminology: “the language . . . seemed really accessible for somebody who's not a med student or a doctor . . . more patient oriented” (P3). In general, variations in medical terminology was a barrier to searching, as participants tended to think in everyday language, but information was often organized and presented using more technical terms, such as use of “influenza” instead of “flu” in the Google Directory. As an alternative to the challenges of selecting an information source in this domain, participants tended to go directly to a known consumer health or government website, commenting, “I'd never research a health topic just using a search engine” (P12). These sites were considered to be more reliable and to provide useful links to other reliable sites.

#### **5.2.5.2 Research**

Research searches tended to focus on the subject content and the format of the information. Participants were very clear on the type of website they were seeking: a reputable site that could serve as a research base: “it was giving an

overview, it was reputable, ...and it looked like it had some other [information], if I wanted to explore further...” (P17). Most of the searches were subject oriented, with proper nouns, i.e., names, used quite often. Participants seemed to think that, “if they're famous, I should put the name in and I should find it ” (p10). However, this approach was problematic if the name was ambiguous or the participant did not spell the name correctly. Some participants searched using keywords designed to limit the results to particular types of information (scientific), formats (journals), or locations (Indian Ocean). When selecting from the hitlist, participants looked for descriptions that indicated that the material would serve as an introduction and starting point, such as “Judaism 101” and “Global Warming brief”. They also tended to prefer websites that took the form of more traditional research sources, such as encyclopedias, newspapers, and books. Websites with no clear textual content were likely to be rejected as “teaser pages” (P32), but too much text was also a problem, as it was difficult to scan. Participants used the links on a website to get a sense of the content, but found this difficult if the meaning of the links was not clear or they were unfamiliar with the topic. Although participants were mainly focused on finding general overview sites, they also valued primary research materials, such as war statistics, population figures, interviews, and historic newspaper articles.

#### **5.2.5.3 Shopping**

Participants were quite specific about the nature of the job: they needed to find a place to purchase a product, and selectively needed to find more information about the product. Those familiar with the product often looked for a store and/or entered the specific name of the product, but this was often not very successful. Products, such as Palm Pilots and the computer game, *Myst*, produced non-relevant (to the user) results due to changes in product names (e.g., Palm V and Exile). Because the task was to purchase, participants considered the need to add “purchasing” to the search statement. Participants scanned the hitlist looking always at the bold titles and often the summary. As one said, “I was looking for something that said, “Hey, here’s a Web cam. Come and get me...I was looking for something that would indicate that they sold the item...”(P25). The indication of price or cost in

the summary was a key cue that the resulting page might enable purchasing. Participants also scanned for cues to indicate the country of origin of websites in order to avoid shipping fees and import taxes, and were frequently frustrated that Canadian sources could not be easily identified. When they could not find the product or a source on the hitlist, they tended to “just put in a store that I knew.”(P10).

The directory was frequently used, presumably because one of the high level categories was “shopping,” but met with mixed success. Participants tended to ‘get lost’ in the hierarchical structure when the category labels were inconsistent and/or incomplete.

Following the selection of a product for purchase, participants often wanted to easily compare products by features and cost. They classified their task into two sub-tasks: one informational to find out about the product and one functional to purchase the product. But switching between these tasks was not easy, as a single site rarely contained both types of information.

#### **5.2.5.4 Travel**

Travels searches were generally quite open and uncritical as to the sources of information. They were characterized by a process in which participants initially were gathering general information about a destination and then were “specifying ... becoming more specific as [they] went on”(P6). Frequently, participants would search using the name of the destination, either alone or with the keywords “travel” or “tourism.” They wanted to find general sites from which they could refine the parameters of their search: “I didn’t know how to approach sailing in Belize, but this immediately gave me a possibility that I could realize” (P12). Those familiar with the destination, tended to be more definite in their queries, often entering specific activities or attractions, e.g. British Museum. Commercial sites, such as Lonely Planet, were sought primarily in the initial stages of the search, and tended to be used by participants with more experience in searching for travel information, who would use them as jumping off points to more specific sites. Concerning the choice between query and directory search, we again suspect that its frequency of use for travel was influenced by a “travel” choice appearing on the Google directory.

Participants were tolerant of large numbers of hits returned: “I figure they are probably arranged in some kind of order according to the search” (P25). A large number of hits was seen as an advantage when comparing and evaluating specific activities or arrangements, e.g. tour packages, hotels, air fares. Hitlists were scanned for the destination name and activities of interest, but this was often frustrating for the participants, as it was difficult for them to discern whether a site was focused on the destination or the activity (heterogeneous task objective). Equally frustrating were those instances when the hitlist description was in English, but the website was predominantly in another language.

Sites were judged on the variety of relevant attributes (e.g. attractions, shopping, restaurants, culture, climate, etc.) they possessed, the links provided, and the presence of maps and images: “You can see what this [the destination] looks like” (P26). Evaluations of authority or credibility only became important when specific factual information was required, e.g. prices, or credentials (scuba instruction).

## **6. Analysis and Discussion**

In this research we found significant differences between domains. Those in Shopping and Travel spent more time exploring within a site, while those in Research and Consumer health spent more time in hitlists. Although this latter finding was not significant, this potentially is a trend. Consumer Health and Research favoured a query approach to the problem, while Shopping and Travel were more inclined to select from the directory categories. We can speculate that the presence of top-level categories semantically related to the assigned task made it easier to use in Shopping, and Travel, and may account for these differences. This may also be rooted in a taxonomic distinction between functional and informational tasks that is not well supported in current search engines [4]. In general Shopping and Travel shared some common characteristics, as did Research and Consumer Health.

Although the quantitative data hinted at differences among the domains, distinguishing elements became apparent from the rich picture presented by verbal protocols. While the intent of each search was to locate information, the approaches taken were distinctly different. Where

Shopping looked for products and stores, Consumer health looked for a trustworthy portal. Research was intent on finding specific genre suited to the domain tasks, and Travel naturally had a geographic orientation. The approaches to query creation, the need for limits in a variety of information types, the way that the hitlist was used and the expectations about what one should find in a hitlist lead us to conclude that, concerning search interfaces, the one-size-fits-all approach may not be the best solution.

## 7. Design Requirements

From our analysis, key issues emerged based on domain that affect the design of the interface, and, may require additional algorithms. Ideally good solutions will be contained in a support layer between the search engine and the interface. For each domain, we have summarized some of the key requirements. These requirements are not exhaustive but serve to illustrate the types of differences that exist, and provide the evidence for more unique support at the interface by domain.

Consumer Health needs:

- a) Hitlist entries must clearly indicate the level (professional vs. lay) and scope (e.g., brief overview or detailed discussion) of information on a site
- b) Hitlist entries must clearly indicate the source (e.g., research centre, pharmaceutical company) of the information
- c) Search interface must explicitly connect medical terms and their everyday equivalents
- d) Search interface must provide the means of limiting search results based on level or scope of information, or source

Research Needs

- a) the ability to specify the level of information (general overview, detailed, scientific)
- b) the ability to specify desired information formats (journal articles, newspapers, statistics, etc.)
- c) a quicker and more effective way to evaluate the content of a website from the hitlist
- d) support for the research process of beginning from a general overview and working towards progressively more detailed information through progressive filtering

Shopping needs:

- a) two types of access: one that enables users to enter queries and one that facilitates product and type selection from a list
- b) the ability to pre-determine functional versus informational needs so that the appropriate results will be supplied
- c) shopping queries must be processed so that product types, brands, product names and stores are distinguished from other types (types of what?).
- d) hitlists must indicate jurisdiction as well as the ability to purchase (and not just look at a catalogue).

Travel needs:

- a) the ability to differentiate between sites that provide more general information (e.g. culture, climate), and those that supply specific travel services (e.g.tour and ticket bookings).
- b) A better mapping between the travel category in a directory and categories representing specific activities. An easy way of establishing the “intersection set.”
- c) Hitlist descriptions that indicate the relative weight of information devoted to keyword terms, e.g. destination vs activity.
- d) Indication in hitlist of whether site offers maps and pictures or is predominantly text website.

## 8. Conclusions

In this study we examined the role of task domain on the search process using a mixed methods research design. We found significant differences among the domains highlighting the need for unique interfaces by domain. The generic search interface is not likely to be as effective for all types of searching. From our work we have identified some of the needs for domain-specific search interfaces.

## Acknowledgements

This work was funded by an NSERC grant to the first author, and by a UBC HSS Grant to the third author.

## About the Authors

Elaine Toms is an Associate Professor in the Faculty of Information Studies, University of Toronto and is currently an IBM CAS Fellow. Her

research interests include the design of tools and methods that improve access to content-rich information systems.

Luanne Freund is a PhD student in the Faculty of Information Studies, University of Toronto and holds an IBM CAS Student Fellowship. Her research interests include improving the query initiation process.

Rick Kopak is an Assistant Professor in the School of Library, Archival and Information Studies at the University of British Columbia. His research interests center on the relationship between information structure and user interaction.

Joan Bartlett is a PhD student in the Faculty of Information Studies, University of Toronto. Her research interests include understanding how bioinformatics researchers access and use data and tools.

## 9. References

- [1] A Student's Guide to Research with the WWW. (<http://www.slu.edu/departments/english/research>).
- [2] Bar-Ilan, J. The Web as an information source on informetrics? A content analysis. *Journal of the American Society for Information Science* 51, 5, (2000), 432-443
- [3] Bates, M. J. Indexing and access for digital libraries and the Internet: Human, database and domain factors. *Journal of the American Society for Information Science*, 49, 13, (1998), 1185-1205.
- [4] Broder, A. A taxonomy of Web search. *SIGIR Forum*, 2002. 36(2).
- [5] Fidel, R. et al. A visit to the information mall: Web searching behavior of high school students. *Journal of American Society of Information Science*, 50, (1999), 24-37,
- [6] Fodness, D. & Murray, B. A typology of tourist information search strategies. *Journal of Travel Research*, 37, (1998), 108-119.
- [7] Fodness, D and Murray, B. A model of tourist information search behavior. *Journal of Travel Research* 37, (1999), 220-230.
- [8] Gursoy, D. Development of a Traveler's Information Search Behavior Model. Unpublished Doctoral Dissertation, Virginia Polytechnic Institute, 2001.
- [9] Hoelscher, C., & Strube, G. Web search behavior of Internet experts and newbies. *Computer Networks*, 33, (2000), 337-346.
- [10] Hsieh-Yee, I. Search tactics of Web users in searching for texts, graphics, known items and subjects: A search simulation study. *The Reference Librarian*, 60, (1998), 61-85.
- [11] Jansen, B. J., & Pooch, UA review of web searching studies and a framework for future research. *Journal of the American Society for Information Science*, 52, 3, (2000), 235-246.
- [12] Kari, J. & Savolainen, R. Web searching in the context of information seeking in everyday life: The cases of civic and spiritual action. A research proposal. Paper presented at Summer school on web searching. 19-21.8.2001. Tampere. Myös <http://www.uta.fi/~csjekar/kari-savolainen.pdf> 27.12.2002.
- [13] Lawrence, S and Giles C.L. Searching the World Wide Web. *Science*, 280, (1998), 98-100.
- [14] Lazonder, A. W., Biemans, H. J. A., & Wopereis, I. G. J. Differences between novice and experienced users in searching information on the World Wide Web. *Journal of the American Society for Information Science*, 51, 6, (2000), 576-581.
- [15] McClung, H.J., Murray, R.D. and Heitlinger, L.A. The Internet as a source for current patient information. *Pediatrics*, 101, 6, (1998), E2.
- [16] Morrison, J. B., Pirolli, P., & Card, S. K. A taxonomic analysis of what World Wide Web activities significantly impact peoples decisions and actions. In J. Jacko, & A. Sears, (Eds). In *Extended Abstracts of the ACM Conference on Human Factors in Computing System*, 163-164, (2001).
- [17] Navarro-Prieto, R., Scaife, M., & Rogers, Y. Cognitive strategies in web searching. *Proceedings of the 5th Conference on Human Factors & the Web*. Gaithersburg, USA, 43-56, (1999).
- [18] Pejtersen, A. M., & Fidel, R. A framework for work centered evaluation and design: A case study of IR on the web - working paper.

MIRA Workshop, Department of Computing Science, University of Glasgow. (1999).

- [19] PEW. Searching for Health Information. in *Pew Internet and American Life Project*, available at <http://www.pewinternet.org/reports/reports.asp?Report=64&Section=ReportLevel1&Field=Level1ID&ID=288>
- [20] Spink, A. and Gunar, O. E-Commerce Web Queries: Excite and Ask Jeeves Study. In *First Monday*, 2001 6(7).
- [21] Spink, A., Jansen, B.J, Wolfram D. & Saracevic, T. From E-Sex to E-Commerce: Web search changes
- [22] Spiteri, L.F. Access to electronic commerce sites on the World Wide Web: an analysis of the effectiveness of six Internet search engines. *Journal of Information Science*, 26, 3 (2000), 173-183.
- [23] Vogt, C.A. & Fesenmaier, D.R. Expanding the functional information search model. *Annals of Tourism Research*, 25, 3 (1998), 551-578.
- [24] Weber, K & Roehl, W. S. Profiling people searching for and purchasing travel products on the World Wide Web. *Journal of Travel Research*, 37 (1999), 291-298.
- [25] Wu, BG, and Li, J. Comparing Web search engine performance in searching consumer health information: evaluation and recommendations. *Bulletin of the Medical Library Association*, 87, 4 (1999), 456-461.



This document was created with Win2PDF available at <http://www.daneprairie.com>.  
The unregistered version of Win2PDF is for evaluation or non-commercial use only.