Evaluating Usability of a Long Query Meta Search Engine

Isak Taksa
Zicklin School of Business
Baruch College
City University of New York
One Bernard Baruch Way
New York, NY 10010, USA
isak_taksa@baruch.cuny.edu

Amanda Spink
Faculty of Information Technology
Queensland University of Technology
Gardens Point Campus
2 George St, GPO Box 2434
Brisbane QLD 4001 Australia
ah.spink@qut.edu.au

Abstract

Usability is an important factor for search engine acceptance. This paper examines usability of a long query meta search engine. The engine was designed to accept and process an unlimited size query expressed in natural language. We briefly review current search engine usability research and then apply some of the common metrics to various tasks of the search and retrieval process beginning with query formulation and concluding with knowledge discovery in relevant search results. We report on users' utilization of many features offered by the engine which enhance the search experience, increase the quality of the search results and improve the usability measurements. Additionally, the implications of this study on the advancement of search engine development are discussed.

1. Introduction

Search engine users are losing the race between the perpetually expanding Web and advances in search engine design. Finding relevant information becomes an arduous and time consuming task. Even if the users know how to express their information needs (IN) in natural language they frequently face a "language" barrier while trying to convert this knowledge into a few exact terms to formulate an adequate search query. Numerous studies confirm users' persistence in using short queries [12, 28]. Another drawback of shorter, unfocused queries is the number of results returned by the search engine. Overwhelmed by the task of sifting through a massive volume of returned search results, users frequently examine just the first page (top 10 results) [19, 26] and continue with a new, reformulated, yet still short query, hoping to find more relevant results. This unstructured approach leads to time waste that is caused by the frequent finding/re-finding and subsequent accepting/discarding of the same documents. To further complicate the problem, very few users are familiar with or use Boolean operators [10] or phrases [32] to improve quality of search results. Additionally, commercial search engines were designed to find information, while leaving users on their own when they try to store and organize the search results [2] or discover new knowledge [15, 34]. In contrast, the long query meta search engine (LQMSE) incorporates a range of novel functions and addresses many of the above concerns. To perform the actual Web search the LQMSE uses Google.

The focus of this research is to measure usability of a long query meta search engine. We present an overview of the current state of usability research and the approach taken in this research. We include a brief description of the functionality of the engine, followed by experiments to measure its usability. We conclude with a discussion of the results and plans for future research.

2. Related usability studies

Researchers in the field of usability have developed a variety of diverse views on what usability is and how to study and measure it. Some, when designing new systems, strictly follow ISO definitions of usability [9] and user-centered design [8], while others apply a user-centered approach to evaluate usability of existing search engines [15]. Some researchers branch out in search of new methodologies and metrics in measuring usability [24], while others try to keep abreast of new and evolving technologies [17]. Many researchers and practitioners agree that context of use should be the driving force in the quest for accurate usability testing and measurements [3, 18]. A recent paper on usability research [7], which examined usability measures from 180 published studies, concluded that choosing an appropriate and acceptable usability measure is a complex and difficult process.

1530-1605/07 $20.00 © 2007 IEEE

The focus of this paper is the usability of a search engine and, therefore, we concentrated on studies that measure usability attributes applicable to various tasks of the information retrieval process. Several studies looked at efficiency and user satisfaction with the query formulation process [22, 33]. Others investigated the effectiveness of this process from a query reformulation point of view. Query reformulation becomes more efficient if the user can see the search terms, which were used in previous search sessions [33]. Providing a larger search field to allow full query view makes the system more useful [21].

Important elements in the users' satisfaction with the information retrieval process are ease of use, continuity of efforts and retention of earlier results. Some researchers built a flow theory to measure user's experience and satisfaction during information seeking activities [20]. Others measure effectiveness of the search engine that provides automated assistance during the search process [11] and detects and eliminates duplicate efforts (submitting same query, examining same search result, etc.) [33].

An important factor of engine usability is its ability to "forgive" errors and allow the user to go back and repeat as many tasks as necessary [30]. Since information retrieval tasks could be performed over a period of time, an engine's ability to retain and reuse earlier results is high on users' satisfaction list [29, 33], especially while conducting multi topic searches.

Because search engines regularly return millions of hits, a search engine that filters search results for user relevance evaluation [13, 24] provides welcome help and improves user satisfaction.

3. Long query meta search engine

Before we proceed with a description of our study and experiments we will briefly discuss conventional information retrieval and how the long query meta search engine changes the established conventions. Information retrieval using a commercial search engine (Google in our discussions and experiments) is a single search event that consists of several tasks. The user starts with information needs, usually expressed at length in natural language. Based on prior experience and domain knowledge the user selects a few search terms to formulate the search query (see Table 1, task 1, Commercial SE) and then submits it to the search engine (task 4).

The search term selection process is inexact and search results could easily become skewed. The user then examines search results returned by the engine (task 5) and then proceeds to either bookmark the relevant site, store its copy on a hard drive, or just discard it (task 6). If search results are not relevant (or if the search engine does not return enough relevant results) the user starts a new search event with some mental recollection of search terms used and sites retrieved in prior search events. Users rarely keep notes of queries and associated search results, so it is a frequent occurrence that similar or identical queries are submitted several times, and the user retrieves, examines, and stores the same results repeatedly. Furthermore, all tasks in this process are manual.

In contrast, the LQMSE redefines all of the above tasks, while also introducing new ones. The search event becomes a search process that automates many functions, thus leaving control of the process in user's hands. It also allows repetition of tasks with the same or different parameters. Instead of converting information needs (IN) into only a few search query terms, the user enters the entire IN description into the search window of the search engine (see Figure 1 below).

Table 1 Search and Retrieval Tasks (Commercial vs. LQMSE) (M-manual task initiated by the user, A-automated engine task)

M/A	Commercial Search Engine Event	TASK #	Long Query Meta Search Engine Process	**M/A**
M	Query formulation	1	Query formulation	A
		2	Query reformulation; Phrase creation	M
		3	Determining filtering criteria	M
M	Submission to SE	4	Specifying control parameters for multiple query formulations; submission to SE	M
M	Search results examinations	5	Search results examinations and ranking	M
M	Storage and management of search results	6	Storage and management of search results	A
		7	Knowledge discovery	A

The engine parses the IN and creates an ordered list of all words (terms) in the IN (excluding "stop" words). The order of words depends on two frequencies: the frequency of the word in the Google collection (number of documents that contain that word) and the frequency of the word in the original IN.

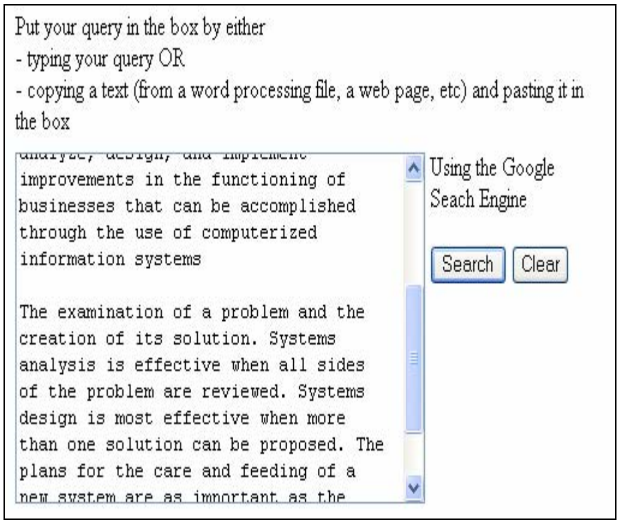

Figure 1 Long Query Search Field

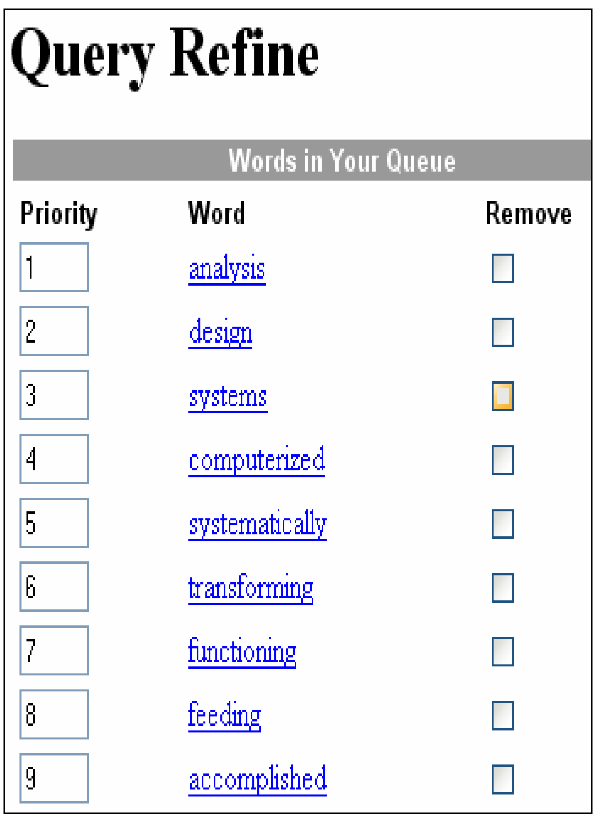

Figure 2 Long query refinement

The ordered list is the initial query formulated by the engine (see Table 1, task 1, LQMSE). The user is presented with this ordered list for editing, i.e. deleting/correcting misspelled words, changing the order of the words or inserting new words. This query reformulation task (task 2) can be repeated any time throughout the search process (see Figure 2 above).

If the user is aware of phrases that are common to the domain, those could be added to the list of search terms submitted to the search engine (see Figure 3 below).

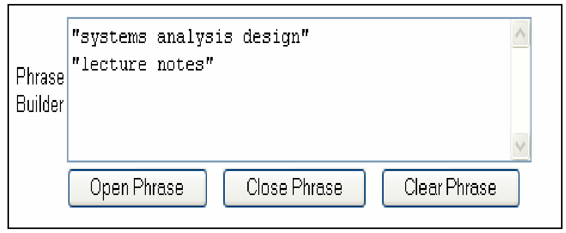

Figure 3 Phrase builder

In task 3, the user can specify the sites that will be filtered-in/out. This excludes the processing and evaluation of unwanted sites while ensuring the inclusion of potentially relevant sites. For example, the user only wants to include (filter-in) sites that offer instructional materials related to Systems Analysis & Design (SA&D) (words like *tutorial, notes, lectures* appearing in text summary or in the URL) and only from US institutions of higher learning (the *.edu/* in its URL) but wants to exclude (filter-out) commercial sites (*.com/* appearing in the URL) that offer the same material (see Figure 4 below).

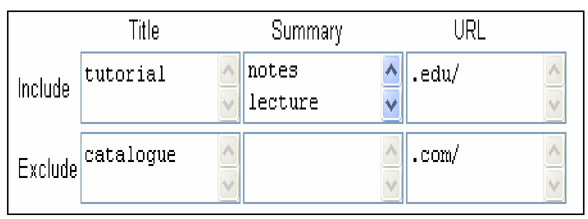

Figure 4 Include/exclude filter settings

The heart of this engine is its ability to create multiple subquery formulations. The user is presented with the option to decide on a number of subqueries (generated from terms in the ordered list, which was created in task 1 and edited in task 2 and submitted to the search engine). In task 4 the user specifies two control parameters ***n***-number of words (top ***n*** words from the ordered list in task 1) that are used to generate conjunctive subqueries and ***r*** – minimum number of words in each subquery. This task could be visualized as spreading a fish net where ***n*** represents the size of the net while ***r*** specifies its mesh. The bigger the ***n***, the wider the net used to "catch" (retrieve) potentially relevant sites. ***r*** on the other hand defines the depth of search (see Figure 5 below). Smaller ***r*** will provide a shallow search, meaning more results are considered for inclusion, larger ***r*** means deeper search where less results are found.

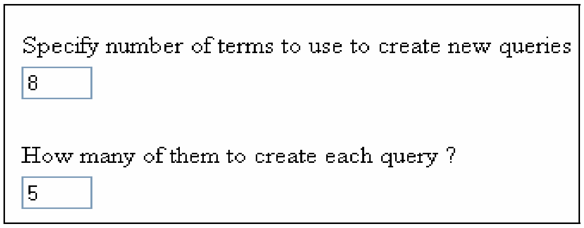

Figure 5 Subquery formulation parameters

The process of subquery formulation is quite straightforward: create various $_{\mathbf{n}}\mathbf{C}_{\mathbf{r}}$ combinations, or simply, create **all** possible search queries consisting of at least **r** terms from the list of **n** terms. The number of such combinations is calculated as follows

$$\sum_{i=r}^{n} nCi = \sum_{i=r}^{n} n!/[(n-i)!*i!] \qquad (1)$$

For example, for **n** = 8 and **r** = 5 the total number of subqueries is 93. Recent research [31] demonstrates that the number of subqueries could be reduced without significant degradation of results.

Once all subqueries are submitted, all search results are filtered according to filter-in/out criteria and combined into a final search result list, which is ordered depending on frequency of each URL in the combined list. The final list is presented to the user for examination (task 5). The user is provided with brief instructions on how to rank and store relevant results.

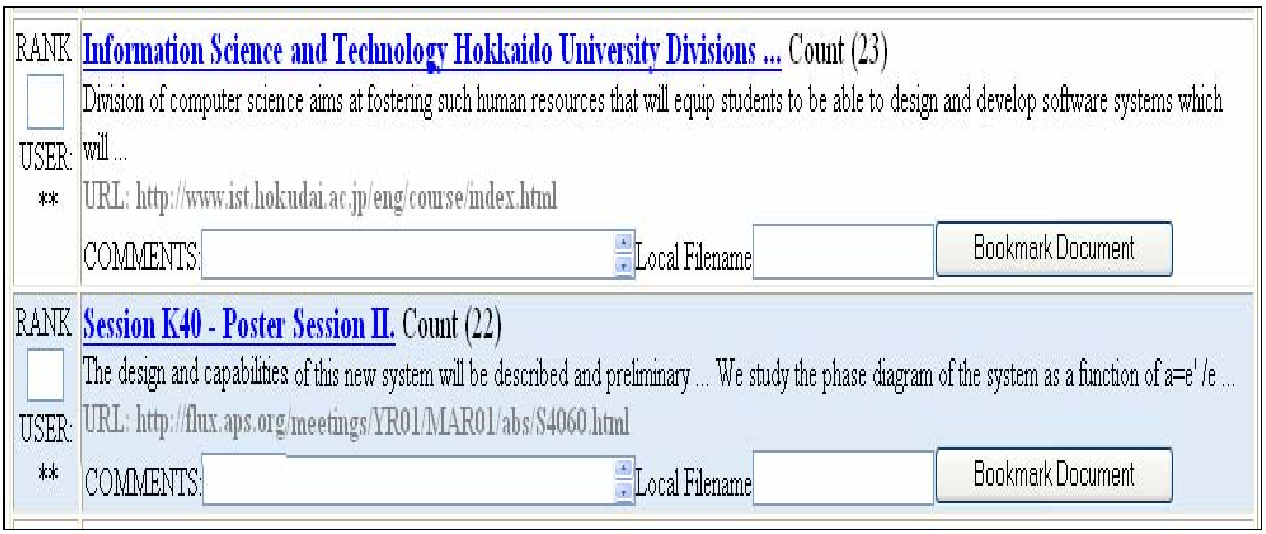

Figure 6 Final results (partial view)

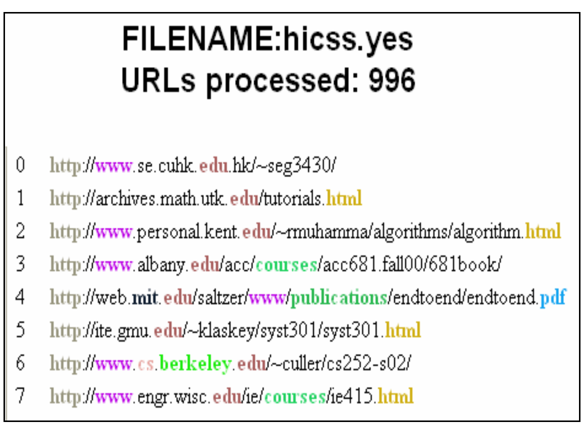

Figure 7 List of relevant results (all URLs)

While examining the site, the user may decide to store or bookmark (see Figure 6 above) the site (by assigning a relevance rank) or just skip it (by leaving the Rank box blank).

The engine stores the ranked results (task 6) and updates two lists of URLs: relevant (accepted) results and non-relevant (rejected) results. URLs in both lists are used to prevent earlier results (relevant and non-relevant) from appearing in subsequent search results, thereby saving the time and effort of processing and examining results more than once.

The final task (task 7) analyzes URLs in both lists for frequent appearance of common words or abbreviations in the *domain name* and/or *path/file* *name* thereby suggesting potential criteria for further filter-in/out choices. Figure 7 above is an example of a file of accepted/relevant sites (called *hicss.yes* and conversely the file for rejected sites is called *hicss.no*). The color-coding and the color legend (see Figure 8) allow for easy visualization of commonalities in a list of URLs.

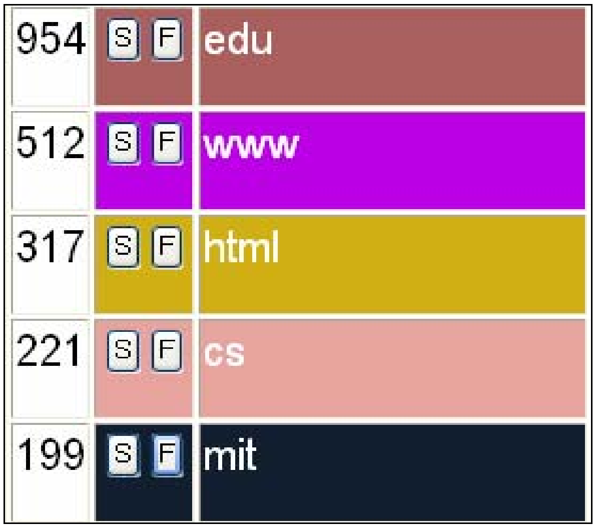

Figure 8 Color legend

For example, observing that there are 199 sites from **mit**, the user may want to investigate other common terms in URL addresses of relevant MIT sites. By clicking button **F** (filter results) the user can observe further URL subdivision for all **mit** sites (see Figure 9 below) in original relevance order with a color legend similar to one in Figure 8. The button **S** provides a list of **mit** sites in alphabetical order.

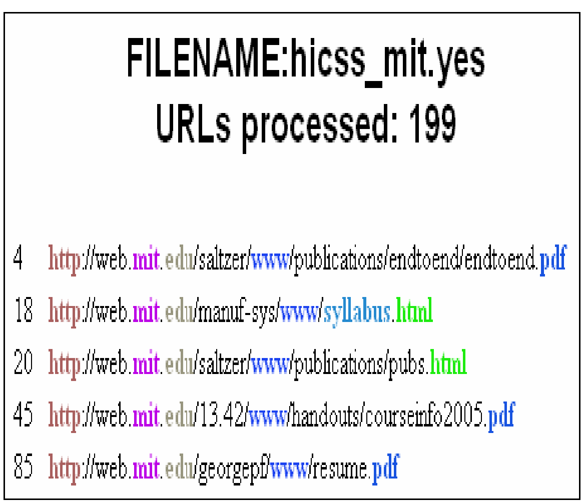

Figure 9 List of relevant results (selected URLs)

4. Research objectives

Usability measures are traditionally classified into three major categories: effectiveness, efficiency, and satisfaction (we follow classification suggested in recent, comprehensive research [7]). Our objectives are to apply some of the conventional measures in each category to usability analysis of the long query search engine. Specifically, in the effectiveness category, *completeness* (user's ability to complete assigned tasks) and *quality of outcome* (relevance of retrieved results); in the efficiency category, *time to complete* (time required to complete the assignment) and *usage patterns* (user's participation in individual retrieval tasks); and in the satisfaction category, *control* (ability to control the outcome) and *learnability and retention* (stress-free learning to learn and hard to forget).

5. Research design

5.1 Participants

The study was conducted on the campus of an international university during the summer session of 2005. Three sections of a multi-section advanced Systems Analysis & Design (SA&D) course were randomly selected. Successful completion of this course required some computer competency. Though research proves that user-competency affects usability results [5, 16], for the purpose of our experiments we assumed that all students performed at the same level. The number of students in each section (group) was approximately the same (30, 32 and 35 students).

5.2 Experiments

A three-part experiment was designed to collect data for usability analysis. In the first week (of a 6-week course) the students received the following two-paragraph description of the concept of SA&D.

> *"The examination of a problem and the creation of its solution. Systems analysis is effective when all sides of the problem are reviewed. Systems design is most effective when more than one solution can be proposed. The plans for the care and feeding of a new system are as important as the problems they solve" [1].*

> *"Systems analysis and design, as performed by systems analysts, seeks to analyze data input or data flow systematically, processing or transforming data, data storage, and*

information output within the context of a particular business. Furthermore, systems analysis and design is used to analyze, design, and implement improvements in the functioning of businesses that can be accomplished through the use of computerized information systems"[14, pp. 6-7].

The first part of the experiment **- Conventional benchmark IR,** was to use the above description to create search queries and to retrieve results relevant to each group's assignment. Each group received an individual assignment to find 20 relevant sites with specific instructions on how to search the Web. Additionally, each participant was asked to rank the relevancy of all selected results using the ordinal relevancy scale (from 1 for non-relevant to 9 for very relevant). We did not use the typical "relevant/non-relevant" discrete scale so as to ensure that students actually open and read each site. To gauge the students' approval of their own search results we asked another student to conduct an impartial relevancy evaluation. The following are the groups' assignments and execution instructions:

Group A: Use as many commercial search engines (SE) and formulate as many queries (Q) as needed to collect a list of sites (URLs) that are similar to the two-paragraph description above (excluding textbooks sellers);

Group B: Use one commercial SE and formulate as many queries as needed to collect a list of sites (URLs) that address the academic issues of SA&D (e.g. scientific papers) discussed in the two-paragraph description above;

Group C: Use one commercial SE and formulate one successful query that will provide 20 distinct educational sites that offer instructional SA&D information (tutorial, lecture notes, etc.).

The second part – **long query meta search engine IR,** which took place in the second week, was to repeat the original assignment while now using LQMSE. Even though students were not familiar with the engine, only **Group C** received an hour of introductory instruction, explaining all seven tasks (as specified in Table 1 above). The other two groups were shown the screen shots for each task for identification purposes only. During the actual experiments, students from **Group B** were allowed to have one-on-one consultation without giving any benefit to other students in the group. No students in any group were allowed to consult with each other.

It is important to mention that this search engine does not have any kind of on-line help, except for Task 6 which instructs the user on how to rank the relevancy of sites the user wants to save for future use. After the experiment, students in all three groups submitted written requests for consultation specifying the task number in question. Observers collected and tabulated these requests and then repeated the introductory tutorial (given initially to Group C) followed by a Question and Answer session.

The third part – **motivated long query meta search engine IR,** which took place in the sixth week, was to repeat the second part of the experiment under new conditions. Students did not see, discuss, or use the new engine. Before the experiments, the students were advised to select only highly relevant search results. Again, no consultation of any kind was allowed during the experiments.

5.3 Data analysis

Results of the first experiment, presented in Table 2, show that students are familiar with commercial search engines, are willing to experiment with search queries and spend time to obtain relevant results.

Table 2 Conventional IR results

Group	# of SE (Min/ Max)	# of Queries (Min/ Max)	# of Terms (Min/ Max)	Time spent (minutes) (Min/ Max)	% of shared results
A	3/5	5/11	3/7	65/90	11
B	1	2/14	2/6	45/90	45
C	1	1	4/9	50/85	67

Before we proceed with the discussion, please review the following Figure 10. It depicts the legend common to all charts shown below.

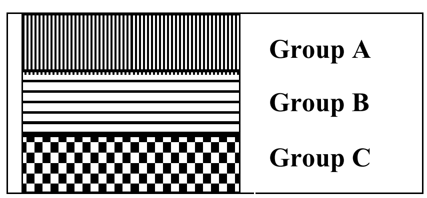

Figure 10 Chart legend

We identified two attributes in the effectiveness category: completeness and quality of outcome. Similar to any commercial search engine, when the user presses the "Search" (or "Go") button to begin the search, it is guaranteed that the search will be completed. Even though LQMSE consists of many tasks, the built-in defaults allowed users in all groups to successfully complete the search.

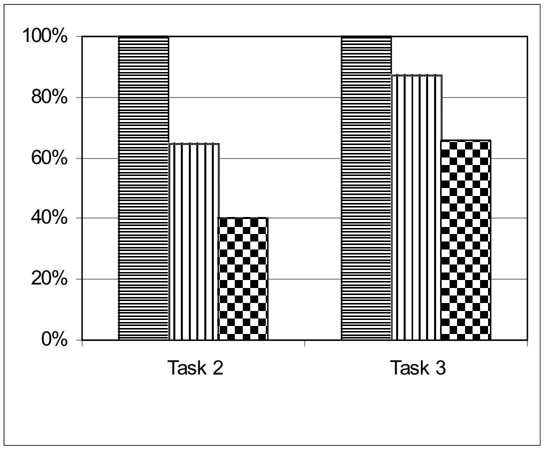

Figure 11 Using the default options to complete the search assignment

While it was an easy assignment for **Group A** (find similar results), students in **Group B** (find scientific papers) had to figure out a way to tweak the engine to retrieve more specific results. They used observers' consultations extensively, but only relied on the query reformulations option, and shied away from the filter option. On the other hand, **Group C** (find instructions material), having benefited from the introductory lecture, experimented with both options.

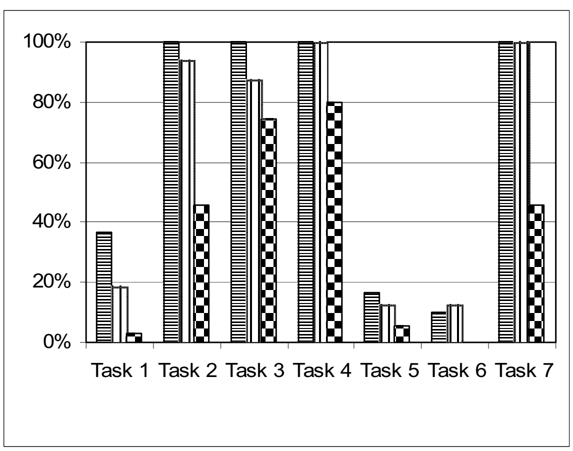

Figure 12 Requests for consultations

Figure 11 above demonstrates the usage of engine defaults to achieve 100% completion of the search assignment. It closely correlates with the students' requests for consultation after completing the second part of the experiment (see Figure 12 below).

To measure quality of outcome, we compared the relevancy results from the first and second parts of the experiments. Table 3 below reflects the fact that impartial evaluation always produces a lower relevancy rank than self-evaluation, and that the discrepancy is consistent among groups and experiments.

Table 3 Relative relevancy ranking Self vs. Impartial

	Part 1	Part 2
Group A	1.75	0.6
Group B	1.6	1.2
Group C	0.9	0.8

In the efficiency category we measured time to complete and usage patterns. Time to complete (we show minimum reported time), as depicted in Figure 13 below, depends on the desired quality of the search results. **Group A** with the easiest assignment completed Part 2 much faster then Part 1. However, when quality requirements were amplified in Part 3, the time to complete increased as well. On the other hand, **Group B** and **Group C** reported longer time to complete Part 3 vs. time reported in Part1. This could be explained by the many iterations of Tasks 2, 3 and 4 performed by the user in order to improve the quality of the final search results. It is important to emphasize, that all tasks (except for Task 5) could be performed autonomously using engine defaults.

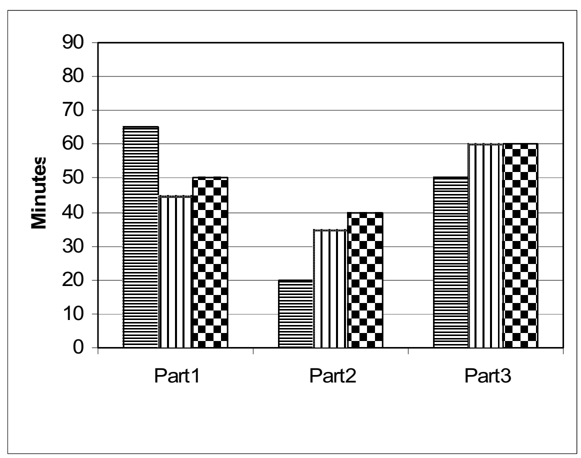

Figure 13 Time to complete

Figure 14 below shows the average number of iterations performed by a student in each group. It is interesting to note that students in **Group C**, who received the most instructional time, were the most frequent users of Task 6, which requires more understanding and potentially improves the final results quicker. On the other hand, Task 4 where the affects on the final results are not obvious, was the least used.

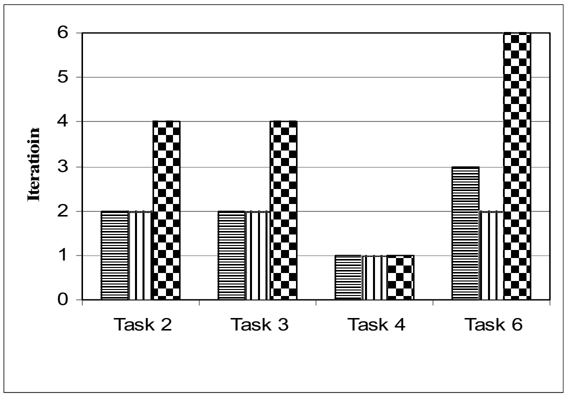

Figure 14 Average iterations per task

In the final satisfaction category we identified two attributes: control and learnability. The first, control, reflects the users' perception of how their action can influence the final search results. Table 4 below shows a dramatic increase in the users' utilization of tasks that control the quality of final search results (when motivated by a higher grade in the course).

Table 4 Utilization of user control (Part 2 vs. Part 3)

Grou	Task 2	Task 3	Task 4	Task 6
A	0/58	0/61	0/32	0/86
B	23/68	14/66	4/29	0/64
C	28/133	18/146	3/40	0/221

The users' ability to retrieve quality results after minimal exposure to a new engine and after a gap of four weeks, is a good measure of learnability and retention (the second attribute of satisfaction). When asked to submit a request for consultation again, after Part 3 was completed, the students' response was largely in the vicinity of 3% for Tasks 4 and 6, and close to 0 for all other tasks.

7. Implication for SE development

Our usability results allow designers of search engines a look at users' behavior, needs and strategy throughout various tasks of the information retrieval process. For example, the *completeness* measure confirmed a need for a well constructed system of defaults to allow even a novice user to complete a search process. On the other hand, the *learnability* measure demonstrated the user's capacity (and motivation) to try new, unexplained and experimental functionality. While this is true for many new technological innovations, the *usage-pattern* measure proved again that the users' participation in an optional task depends on their clear understanding of the costs and benefits associated with the task. The *control* measure confirmed that users armed with this understanding will attempt to manipulate search results and manage the flow of the process.

Our usability results suggest that search engine designers should introduce more functionality that assist the user during the lengthy, comprehensive, and often imprecise information retrieval process. Developers, through on-line or context-sensitive help, should make clear the costs and benefits of using, misusing or not using available functionality. Finally, users should be empowered to control search, retrieval and management of search results throughout the information retrieval process.

8. Conclusion and further research

After examining numerous usability measures, we applied selected measures to evaluate the usability of a long query meta search engine. Our initial results demonstrate that it is possible to select a generalized set of usability measures to evaluate a specialized search engine. Furthermore, we captured results that are significant to the design and development of new search engines.

Our research is still in progress and we plan to expand it in several directions. Additional usability measures will be explored. The experiments will be designed to separately examine objective and subjective usability categories.

In order to obtain more generalized results the meta search engine will use additional underlying search engines. Finally, the research will benefit if users are divided into groups according to search experience and domain of interests.

9. Acknowledgements

This work was supported by a grant from The City University of New York PSC-CUNY Research Award.

10. References

[1] Answers.com http://www.answers.com/topic/systems-design

[2] A. Aula, N. Jhaveri and M. Käki, "Information search and re-access strategies of experienced web users," In *Proceedings of the 14th international conference on World Wide Web*, Japan, pp. 583–592, 2005

[3] C. Cool and A. Spink, "Issues of Context in Information Retrieval (IR): an Introduction to Special Issue," *Information Processing and Management*, Vol. 38(5), pp. 605–611, 2002

[4] A. Dillon, "Beyond usability: process, outcome and affect in human computer interactions," *Canadian Journal of Library and Information Science*, Vol. 26(4), pp. 57–69, 2002

[5] L. Faulkner and D. Wick, "Cross-user analysis: Benefits of skill level comparison in usability testing," *Interacting with Computers*, Vol. 17(6), pp. 773–786, 2005

[6] H. Greisdorf and A. Spink, "Median Measure: An Approach to IR System Evaluation," *Information Processing and Management*, Vol. 37(6), pp. 843–857, 2001

[7] K. Hornbæk, "Current practice in measuring usability: Challenges to usability studies and research," *International Journal of Human-Computer Studies*, Vol. 64(2), pp. 79–102, 2006

[8] ISO 13407:1999, Human-centered design processes for interactive systems
http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=21197

[9] ISO 9241-11:1998, Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability
http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=16883&ICS1=13&ICS2=180&ICS3=

[10] B. J. Jansen and C. M. Eastman, "The Effects of Search Engines and Query Operators on Top Ranked Results," In *Proceedings of International Conference on Information Technology: Computers and Communications*, Las Vegas, 2003, pp. 135–139

[11] B. J. Jansen and M. D. McNeese, "Evaluating the Effectiveness of and Patterns of Interactions With Automated Searching Assistance, *Journal Of The American Society For Information Science And Technology*, Vol. 56(14), pp. 1480–1503, 2005

[12] B. J. Jansen and A. Spink, "How are we searching the World Wide Web? A comparison of nine search engine transaction logs," *Information Processing and Management*, Vol. 42(1), pp. 248–263, 2006

[13] M. Käki and A. Aula, "Findex: improving search result use through automatic filtering categories," *Interacting with Computers,* Vol. 17(2), pp. 187-206, March 2005

[14] K. Kendall and J. Kendal, *Systems Analysis and Design*., Upper Sadle River, Pearson/PrenticeHall, 2005

[15] C-H. Li and C-C. Kit, "Web Structure Mining for Usability Analysis" In *Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence (WI'05)*, pp. 309–312, 2005

[16] G. Meiselwitz and G. Trajkovski, "Effects of Computer Competency on Usability and Learning Experience in Online Learning Environments," In *Proceedings of Seventh ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD'06),* pp. 339–342, 2006

[17] A. Monk, "Noddy's guide to usability," *Interfaces*, Vol. 50, pp.31–33, 2002

[18] J. Nielsen, Usability engineering. San Diego, Academic Press, 1993

[19] S. Ozmutlu, A. Spink and H. C. Ozmutlu, "A Day in the Life of Web Searching: an Exploratory Study," *Information Processing and Management*, Vol. 40(2), pp. 319–345, 2004

[20] S. Pace, "A grounded theory of the flow experiences of Web users", *International Journal of Human-Computer Studies,* Vol. 60(3), pp.327–363, March 2004

[21] D. E. Rose, "Reconciling Information-Seeking Behavior with Search User Interfaces for the Web," *Journal of the American Society for Information Science and Technology,* Vol. 57(6), pp. 797–799, 2006

[22] B. Schmitt and S. Oberländer, "Evaluating and Enhancing Meta-Search Performance in Digital Libraries," In *Proceedings of the 3rd International Conference on Web Information Systems Engineering (WISE.02)*, pp.93–104, 2000

[23] H. Sharma and B. J. Jansen, "Automated evaluation of search engine performance via implicit user feedback," In *Proceedings of the 28th annual international ACM SIGIR conference on research and development in information retrieval*, pp. 649–650, 2005

[24] M. Shepherd, C. Watters and A. Marath, "Adaptive User Modeling for Filtering Electronic News," In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS'02),* Hawaii, 2002, Volume 4, pp. 102b

[25] A. Spink, "A User Centered Approach to evaluating Human Interaction with Web Search Engines: an Exploratory Study," *Information Processing and Management*, Vol. 38(3), pp. 401–426, 2002

[26] A. Spink and B. J. Jansen, *Web search, public searching of the web*. New York, Kluwer, 2004

[27] A. Spink, M. Park, B. J. Jansen and J. Pedersen, "Multitasking During Web Search Sessions," *Information Processing and Management*, Vol. 42(1), pp. 264–275, 2006

[28] A. Spink, D. Wolfram, B. J. Jansen and T. Saracevic, "The public and their queries," *Journal of the American Society for Information Science and Technology*, Vol. 52(3), pp. 226–234, 2001

[29] T. Sumner and M. Dawe, "Looking at Digital Library Usability from a Reuse Perspective," In *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries*, pp. 416–425, 2001

[30] A. Sutcliffe, "Assessing the Reliability of Heuristic Evaluation for Website Attractiveness and Usability," In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS'02),* Vol. 5, pp. 137, 2002

[31] I. Taksa, "Predicting the Cumulative Effect of Multiple Query Formulations," In *Proceedings of the IEEE International Conference on Information Technology: Coding and Computing,* Vol. II, pp. 491–496, April 2005

[32] H. Topi and W. Lucas, "Searching the Web: operator assistance required," *Information Processing & Management*, Vol. 41(2), pp. 383–403, March 2005

[33] B. M. Wildemuth, "Evidence-based practice in search interface design," *Journal of the American Society for Information Science and Technology*, Vol. 57(6), pp. 825–828, 2006

[34] H. Wu, M. Gordon, K. DeMaagd and W. Fau, "Mining web navigations for intelligence," *Decision Support Systems*, Vol. 41(3), pp. 574–591, March 2006