Using Eye Tracking To Evaluate Alternative Search output Interfaces

Rachana S. Rele
Industrial Engineering
Clemson University, Clemson, SC 29634
rrele@clemson.edu

Andrew T. Duchowski
Computer Science
Clemson University, Clemson, SC 29634
duchowski@acm.org

ABSTRACT

An increasing number of web pages are indexed by various search engines, thus making it challenging to present this information to web searchers in a more effective and efficient way. Surveys have shown that 75% of users get frustrated with search engines, and only 21% of users reported to have found what they were looking for every time. In addition to problems encountered during query formation, inability to find relevant set of results can be attributed to the interface design of the search results page.

This paper presents results from a study that used sixteen subject's to evaluate two search result interfaces, list interface commonly seen on many search engines, and a tabular interface as an alternative to the list interface. Previous research has devised different display techniques for presenting such information to the users. One such study has shown that a tabular interface for presenting search results is both objectively and subjectively better than the conventional list interface used by most search engines. The two tasks used in the study were named based on the taxonomic research by Broder et al [1]; navigational task that would require users to look for a particular field on the results page, and an information seeking task that will not create any bias to scan a particular field. In the current research eye tracking was used to evaluate alternative interfaces. Quantitative comparisons of two interfaces are made on performance metrics, such as time, and errors; process metrics, such as fixation durations, number of fixation, eye movement transitions. Subjective data was collected through questionnaires. This study attempts to uncover the underlying ocular behavior, and relate this with their performance on two types of interfaces, and two types of tasks. The results provide some insights into importance of particular results categories such as title, summary, and URL of the interface while searching.

Author Keywords
Eye Movements, Search Results, Tabular Interface, List Interface.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Information search has become one of the most frequent activities on the net according to the RealNames research [2] conducted in 2000; web users spend 70% or more of their time searching online. Forrester report [3] in 2000 stated that search engines are the top way consumers find new web sites online, used by 73.4% of those surveyed. In a "Search Rage" study conducted by WebTop in 2000 nearly 75% of the respondents reported frustration of some significant degree when asked *"How frustrating do you find getting irrelevant information when web searching?"* In a separate question, *"Do you feel that Web searching could be more efficient?"* the vast majority said yes: 86 %. Only 9 percent felt things were fine as they are. Searchers generally only visit the first three web sites listed in search results, and one out of five visits will last for a minute or less, based on an analysis of 450,000 queries run by AllTheWeb [4] in a 24 hour period. Additionally, this study mentions that, more than half of all searchers will visit only one site in the top results and more than 80% will stop after visiting three. Only 19% will go to the second page of results and fewer than 10% go to the third page.

These statistics imply that for search engines to be efficient, either need to train their users with engine specific Boolean logic for query formation or present the results in a way that would help users to scan efficiently for the most efficient or relevant ones. In a recent study, Bandos et al. [5] devised search interfaces to help users with advanced operators to formulate their search queries. Though this strategy will increase the number of relevant results found, it will also increase the task time, hence further validation is required to know whether users would actually follow the instructions to employ the Boolean logic, or search for the well presented results on the interface.

Alternate approaches can be devised to increase efficiency of search engines' output, by designing an interface layout that affords easy scanning of search results, and does not require users to learn the Boolean logic. Majority search engines use a standard format, that is, list format for displaying results. However, research needs to be done on the efficiency of this linear presentation of data, and usefulness of various fields to the users based on the tasks they perform, and scanning strategies they adopt. For example, while searching for Michael Jordan's home page, users need not go through all the categories of the presented result such as the title, summary, URL, date and field. Due to linear presentation in list format this task may not be achieved in the most efficient way, requiring users to read all the information present in the result. By designing interfaces that allow users to consider or discard particular category of the result can potentially reduce search time and thus increase efficiency of search.

The research reported in this paper evaluates two search results interfaces namely, list and tabular, for two types of task. The ocular behavior of the users on the first page of the search results is compared and used to evaluate the interface layouts. Tobii 1750 binocular eye tracker is used to capture eye movements of the participants. The primary hypothesis of this research is that tabular interface will increase efficiency and accuracy in parsing search result through spatial grouping of results into distinct category columns. Additionally, the scanning strategies adopted on the two interfaces will differ.

BACKGROUND

Eye-Tracking and Usability

Eye tracking is an underutilized tool in web usability studies, primarily due to the large amount of data it produces, and difficulties in data reduction and analysis. Moreover, the eye trackers in the past have been very intrusive and hence the most unnatural way of evaluating interfaces. However, this reason for not using eye trackers in web usability studies is beginning to vanish with the advent of fourth generation non-intrusive eye trackers such as Tobii eye trackers.

Earlier work by Goldberg et al. [6] in eye tracking and usability has shown that the results obtained from eye movements can greatly enhance the understanding of user strategies while interacting with computer interfaces, and this can improve interface evaluations. Previous eye tracking studies on the web are mainly related to media research. One such study by Schiesl et al [7] has shown dissociation between the eye tracking data and the self reported subjective data in a usability study of a news website, thus demonstrating how eye tracking can be effectively used to know users focus of attention, relying less on subjective data from users, and arriving at the problem areas of the interface. Additionally, the Poynter study [8] focused on the viewing/reading behavior of the web news readers. The results from the conventional methods such as focus groups, usability testing, and log analysis, when combined with eye tracking data can help in designing news sites. The results from the study can help develop guidelines for designing news websites. In yet another study in eye tracking by Goldberg et al. [9] addressed specific design issues of features such as portlets, and the objects within the portlets, for a prototype web portal application, the results from measuring eye movements provided a basis to claim that search did not become directed as the sequence of the pages viewed increased.

The aforementioned studies have demonstrated the use of eye tracking in evaluation of user interfaces through a more user centered way than the usual usability metrics of performance.

Search Output Interface

In previous research, Resnick et al. [10] explained various search strategies used by search engine user's to evaluate the search results on two interfaces. These strategies partially depend on the output interface. One of the tasks use to evaluate the interfaces was an information search task. The two output formats used are (a) List-used by Google, and (b) Tabular-presented results in columns corresponding to each element of the result in the list format. The results show that with list format 67% participants used self terminating search (selecting the first result that meets minimum match criterion) whereas with the tabular format, only 50% of the searcher chose to terminate their search when satisfactory match was identified. The self reported subjective data shows that for the tabular layout, users scanned only one field for all options until they found one that met their match criterion, and this layout was also the preferred among the two interfaces. However, eye tracking research can be used to further validate these results. In a study conducted by Dumais et al. [11] seven search result interfaces were created and tested with users. These interfaces were list interfaces, category interfaces, and a combination of the two in which the search results were arranged in various preexisting categories on-the-fly by automatic text classifiers. The results have shown that the category interfaces were faster than the list interfaces. The best performance by users was achieved in the interface having categories along with summaries of individual result. The performance is attributed to the information grouping that category interface employs, thus enabling users to completely discard certain categories while considering the ones which best represent the context of use.

In a study by Resnick et al. [12] described the best practices that can be used to present the search results to the users. This paper had four studies that described search output interfaces in different context. For example, one study discussed the search output interface for a site specific search engine, whereas another considered designing different search results layout according to the user profiles.

All the above mentioned studies have looked at the search output interface both objectively and subjectively. However, the current research discusses performance achieved by the users and the underlying processes evident from the eye movement's data.

To best of our knowledge there are four eye-tracking studies that have attempted to explain how users browse through the presented results of their search query. Granka et al. [13] explored how users viewed the presented abstracts and how they select links for further explorations. This study also gave a relationship between the number of results viewed above and below a selected document and the rank of these viewed results in the results list. Pan et al. [14] evaluated various factors that contribute to viewing behavior on the web. The results specific to search sites indicate that the mean fixation duration on the 1st and the 2nd page of the search results remained fairly constant, while users spend less time fixating on the 2nd page of the search results. Additionally, saccades rate reduced on the 2nd page of the search results page. Thus there are contradicting results from two measurements, such as fixation rate indicated that 2nd page of the search results is easier to scan while saccade rate showed that task difficulty increased on the 2nd page. Salvogarvi et al. [15] in their eye tracking research used only three participants, and measured the pupil dilation of the searchers. The results show that pupil dilation increased while viewing relevant abstracts. However, due to few participants used in this study, the results need validation with a larger set of users. In a recent study by Klockner et al. [16] investigated whether users employed the breadth first or depth first strategy in scanning the results list. In the breadth-first strategy users scan all the results before opening any document, whereas, in the depth first strategy users examines each result in the list and decides immediately whether to open the document in question. The results indicate that 65% users employed a depth first strategy, and a minority of 15% showed the depth-first strategy. A partial breadth-depth first strategy was used by remaining 20% of the users. This depth first strategy is similar to self-terminating strategy mentioned by Resnick et al. [10].

As 90% users view only the first page of search results, indicated in the study by Jansen et al. [17], Silverstein, [18], this study will evaluate the interface layouts based on the viewing behavior, and performance on first ten results or the first results page of the search query. The independent variables used are type of interface and type of tasks. The dependent variable are tome for tasks, errors or number of wrong choices, eye fixation duration, number of fixations in each category of the results, probability of making transitions in the same category of results.

METHODOLOGY

Subjects

Sixteen undergraduate and graduate students (6 Females, 10 Males) at Clemson University performed four tasks, two information search tasks, and two navigation tasks. The age range of the participants was 20.5 to 29 years with a mean of 24.9 years. All the users had a minimum of 5 years internet experience, and searching information was one of their daily internet activities, Google being their primary search engine.

Apparatus

The study used a Tobii 1750 binocular eye tracker integrated with 17" TFT display having a maximum resolution of 1280 X 1024. This is a non intrusive eye tracker which does not require the subject to wear a helmet or any other markers. Figure 1 shows the Tobii eye tracker ised for this experiment. The eye tracker has a tracking rate or the frame rate of 50 Hz, and looks like a normal computer display with cameras and illuminators hidden behind filters. Hence eye tracking becomes nearly invisible to the user. The Tobii hardware consists of a camera with a high resolution, and large field of view used to capture images on the subjects eyes. The NIRLED's (Near Infra Red Light Emitting Diodes) are used to generate even lighting and reflection patterns of the subject's eyes. The accuracy of gaze estimation is 1cm at 50 cm viewing distance, and average accuracy of 0.5 degrees. The latency for the eye tracker is 25-35 ms, and has +/- 3 ms timestamp accuracy. The system is fairly tolerant to large head motions. Tobii uses ClearView software that facilitates preparation, recording and analyses the data collected. The ClearView gives the visualization of the scanpath in terms of the gaze data points, and hotspot plots that use color coding to indicate fixation durations.

The viewing Tobii screen subtended a visual angle of 28 degrees horizontally and vertically at the participants' eyes. The textual stimulus had two fonts a 12pt font for the Title category of the results, and a 10 pt font for rest of the text on the interface. The above mentioned fonts subtended visual angles of 0.003472 degrees and 0.00347 degrees respectively.

Stimuli

The stimuli presented to the subjects are shown in the following figure. Figure 2 shows the list interface for an information task, and Figure 2 shows the tabular interface for the same task.

Figure 1: Tobii 1750 eye tracker

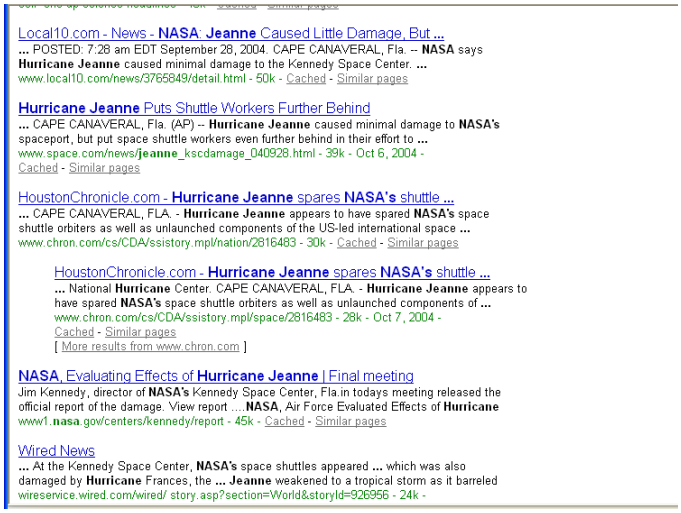

Figure 2 List interface

Figure 3 Tabular interface

Search results for the tasks were obtained using Google search engine. Two information search tasks such as searching for NASA's report on damage to its space center at Florida, and searching for a moon skating store that sells skating accessories were considered of similar difficulty. Two navigation tasks, such as finding the home page for Michael Murray, the mathematician, and finding homepage for university that integrates Stirling engines in its curriculum were considered of similar difficulty. This gave rise to eight different trials (4 tasks X 2 interfaces), four performed by each participant. Thus any participant performed a total of four tasks, two information tasks, one on list and another on tabular interface, and two navigation tasks, one on list interface and the other on tabular interface. The correct result position was randomized as the pilot tests indicated that participants were looking for a pattern of the correct result. The entire experiment was simulated to look like a web search activity, to present the same set of results to all the participants.

Experimental Design

A two factorial design was used with two factors being interface type (List and Tabular interfaces) and task type (information and navigation tasks), each having two levels. The order in which participants viewed the tasks was counterbalanced.

Information tasks

1. Finding information on the report published by NASA, detailing Hurricane Jeanne's damage to its space center at Florida.

A: Tabular interface

B: List interface

2. Finding information about the moon skating store that sells skating accessories.

C: Tabular interface

D: List interface

Events	Subjects			
	1	2	…	16
Familiarization	Tabular	Tabular	…	Tabular
Task 1	A	B	…	B
Questionnaire	✓	✓	…	✓
Task 2	D	C	…	C
Questionnaire	✓	✓		✓
Task 3	E	F	…	E
Questionnaire	✓	✓	…	✓
Task 4	H	G	…	H
Questionnaire	✓	✓	…	✓
Post-test Questionnaire	✓	✓	…	✓

Table 1: Experimental Design

Navigation Tasks

1. Find home page of Michael Murray, the mathematician.

E: Tabular interface

F: List interface

2. Find home page of a University that uses Stirling engines in its curriculum.

G: Tabular interface

H: List interface

Presentation of tasks and interfaces was counterbalanced across subjects.

Procedure

Subjects were screened for their experience with the internet (minimum 5 years). After seeking informed consent, participants were familiarized with the tabular interface in an un-paced fashion. Instructions for the study followed the familiarization phase. The tasks were presented to them on the simulated search query input interface, with a query input box, and a search button. The query was presented and the keywords for the tasks were entered by the experimenter to make the results comparable across participants. There was only one result which was considered to be the right one. Hence, if the participant clicked on a wrong result, a go back page would appear asking the participants to start searching for the correct result. The trial was terminated when the participant found the correct result. Each participant encountered four such trials, which are four tasks, two tasks on list interface and equivalent tasks on the tabular interface. At the end of each tasks participants answered questions specific to the task, and at the end of the experiment, they were asked a few questions regarding their preference, and satisfaction with the two interfaces. One experimental session lasted for abit 30 minutes, including calibration for the eye tracking purpose.

Data was collected on response variables such as:

Performance: Time, and Errors

Process: Fixation Durations, Number of fixations, Transitions to and from different AOI's, percentage of fixations in different categories of the results.

Subjective: Perceived time for task completion, perceived accuracy, ease of finding information on the two interfaces, preference, and overall satisfaction.

RESULTS

Time and errors: The time for tasks did not significantly differ for the two interfaces (F=2.34, p>0.05), and for the two tasks (F=0.77, p>0.05). Moreover there was no significant (F=0.66, p>0.05) interaction between the type of task and interface type for the time values.

The errors or the wrong number of results chosen were not significantly different for tabular and the list interface (F=0.16, p>0.05). Additionally, the types of tasks did not differ significantly on the number of errors made in choosing the correct result (F=0.03, p>0.05); there was no significant interaction effect between interface type and task type (F= 0.27, p>0.05).

Eye movements: The probability of making a transition in the same category of the result was significantly different (F=111.32 p<0.001) for list (0.16) and tabular interface (0.58). Figure 4 shows the mean probability of same category transitions for list, and tabular interfaces. However, this probability was not significantly different (F=0.16, p>0.05) for type of task, and no significant interaction (F=0.03, p>0.05) was found between the interface type and the type of task.

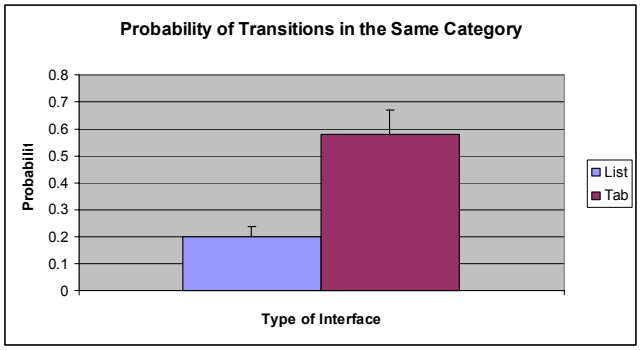

Figure 4: Mean probability of transitions made in the same category for the two interfaces.

The mean fixation duration for the two interfaces did not significantly differ (F= 1.99, p>0.05). The same results were seen for the type of tasks showing no significant difference (F=0.25, p>0.05) in the mean fixation duration. Additionally, there was no significant interaction between the type of task and the interface (F=0.30, p>0.05). Figure 5 shows the mean fixation durations for two types of tasks on two interfaces.

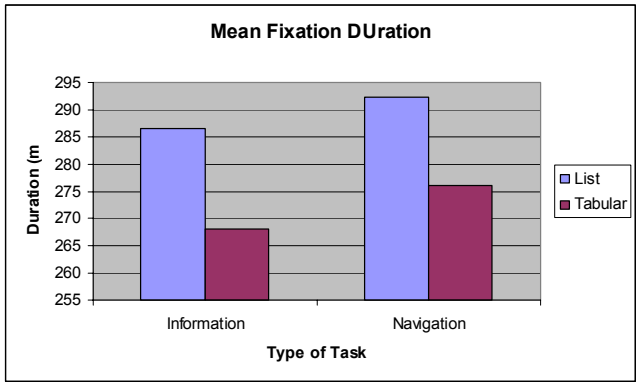

Figure 5: Mean fixation durations on the two interfaces, for two types of tasks.

The number of fixations in the title AOI's of the two interfaces did not significantly differ (F= 0.55, p>0.05). However, the number of fixations in the summary AOI's significantly differed (F= 7.4, p<0.01) for the two types of tasks, navigation tasks requiring more number of fixations (LSmeans=42.68) than the information task

(LSmeans=24.15). The number of fixations in the summary AOI's did not differ significantly (F= 0.14, p>0.05) for the two interfaces. The number of fixations in the URL AOI's did not significantly differ (F= 0.20, p>0.05) for the two types of tasks. However, there was a significant difference (F=11.55, p<0.01) between the numbers of fixations falling in the URL AOI's for list interface and the tabular interface.

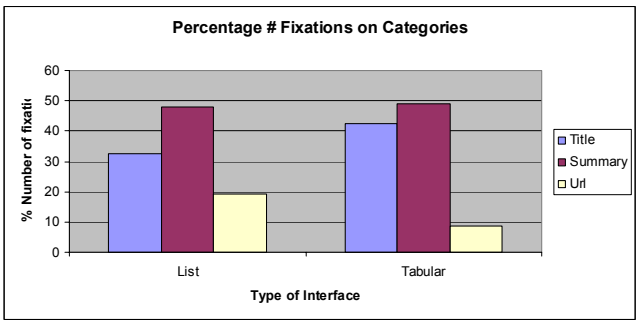

Figure 6: Percentage fixations on different categories for the list and tabular interfaces.

Subjective questionnaire results: Wilcoxon ranked sum, and signed rank test were used to analyze the subjective data gathered.

There was no significant difference (F=2.0, p>0.05) in the preference for the two interfaces.

Overall satisfaction with the kind of interface did not significantly differ (F=0.75, p>0.05) for the list and tabular interfaces.

Perceived time of task completion on list interface and tabular interface did not significantly differ (F=0.87, p>0.05)

Perceived accuracy for the two interfaces did not differ (F=2.14, p>0.05)

Ease of finding information on the list interface did not significantly better from that on the tabular interface (t=1.16, p>0.05)

DISCUSSION

Time taken on the two interfaces did not significantly differ, although time for tabular interface (mean=58.6) was marginally more than time on the list interface (mean=42.1). This can be attributed to the practice effect of viewing results on search engines such as Google. The un-paced familiarization that was given to the participants on the tabular interface may not have been sufficient to equalize the tabular and list interface in terms of user expertise. Duchowski [19] has shown that familiarity of the visual display influences fixation duration, and since fixations contribute to 90% of the viewing time, this result can be attributed to unequal familiarity with the two displays.

The errors of clicking on the wrong results did not significantly differ for the two interfaces, and for the two types of tasks. This indicates that users did not change their clicking behavior depending on the interface. The results from errors can be supplemented with the viewing behavior before clicking such as the transitions made can provide some insight into the strategy used for a particular interface.

The probability of making a transition to the same category was significantly more for the tabular interface than for the list interface. This indicates that users selectively attended to a particular category due to the vertical arrangement of data, hence showing tendencies to move within columns than between columns. Thus users could prioritize categories using the tabular interface. This result is different from that obtained by Goldberg et al., [9] here they found that different portlet visits were more likely than the same portlet visits. However, the stimulus used in this study differs from that of the above mentioned study. Moreover, there could be more than one portlet in a particular column. This result indicates that tabular interface induces a same column bias to the peripheral visual system, which then decides to make the next eye movement. However, no data was collected to know the relative use of the peripheral visual system in the two interfaces. No significant difference was found the probability of making same category transitions for the type of task; this means that scanning strategy remained the same for two different types of tasks. This indicates that regardless of the task context users adopt the same strategy, and there is a major influence of the type of interface they view.

The mean fixation durations are content independent measures [6] and hence any difference in this metric can be reliably attributed to the interface design. However, there was no significant difference in the mean fixation durations for the two interfaces, although mean fixation durations increased for list interface (289ms) than tabular interface (271ms). The list interface may have induced more cognitive effort than the tabular interface.

The total number of fixations did not significantly differ between the list and the tabular interface, indicating that the users may have processes approximately same number of components on the two interfaces. However, number of fixations does not give any depth of processing of the interface components by the users.

The number of fixations in the title category of the results did not significantly differ between the two interfaces, though there are more fixations for the tabular interface (30.2) than the list interface (21.75). This can be attributed to the vertical presentation of titles in a column that guides them to travel vertically. Users fixated more percentage of times on the title category in the tabular interface (42.58%) than that in the list interface (32.58%)

The number of fixations on the summary category of the results for navigation tasks was significantly more than for the information task, though there was no effect of

interaction between the type of task and the interface viewed. There were same percentage number of fixations on the summary category for the list (48.12%) and tabular interface (48.92%). The navigation task of finding the homepage of a university that incorporates Stirling engines in its curriculum may have been artificially difficult; this was evident from the subjective data that was collected. This difficulty was evident from the user comments that they were searching for information on the curriculum and ignored the university URL information, which was the correct answer to the query. This may have lead to more amount of reading in the summary category, to confirm their choice.

The two hotspot plots are shown below for the task of finding information on NASA's report on Hurricane Jeanne. The correct result for the query was on 8th position. Figures 7, suggests that in tabular interface users are evaluating all the results, and fixating on the correct result more prominently, unlike Figure 8, in which users using the list interface do not have comparatively less fixations on the target result than the non target ones.

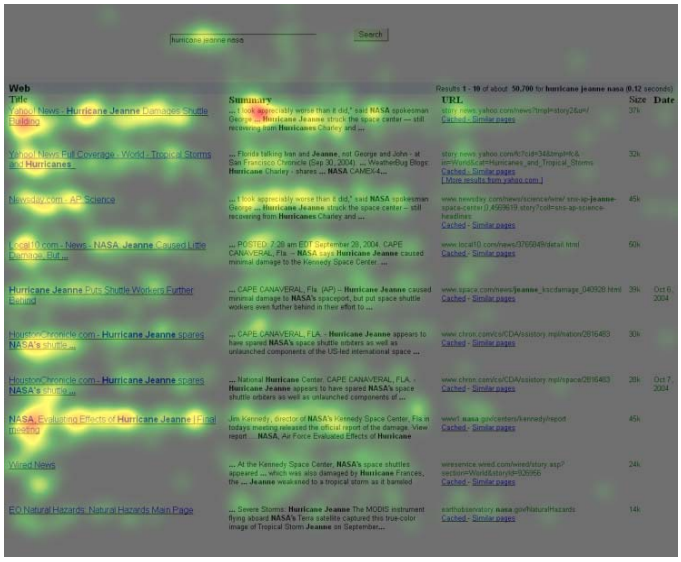

Figure 7: Hotspot plot for an information task on the tabular interface.

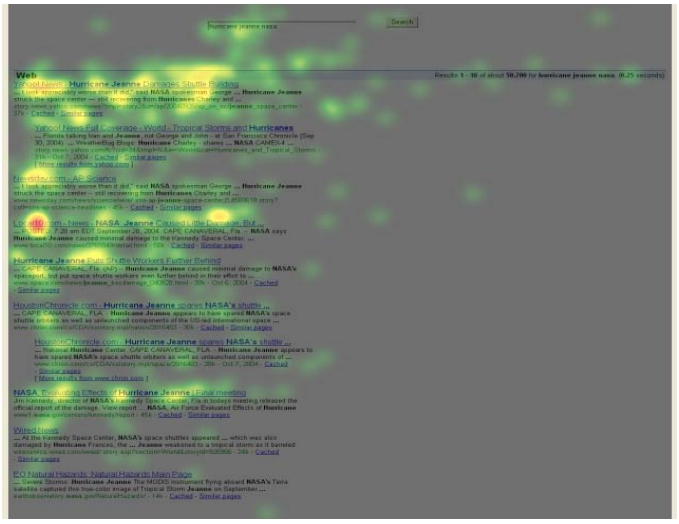

Figure 8: Hotspot plot for information task on the list interface.

Number of fixations in the URL category of results significantly increased for list interface than for tabular interface, this may be due to viewing of the results linearly on the list interface that led to reading everything that was encountered linearly. The percentage number of fixations on the URL category for the list interface was 19.28% as compared to 8.49 % fixations in the URL category on the tabular interface. This denotes that users did not discard the URL category on the list interface, but did so in the tabular interface. This signifies that users do not attend to all the information in the same way. They place weights to different elements of the results and thus selectively scanning these according to the weights, was made possible by the tabular interface.

The subjective data indicates that preference for the two interfaces did not significantly differ. However, Resnick et al found that tabular interface was a preferred interface. This difference in the preference may be due to the different tasks that were chosen by these two studies. The finding for preference of the interface though not significant, the subjective rating for the tabular interface for a minority of users (37.5%) is very high. Hence, this suggests that 37.5% users can be benefited from the tabular presentation of data. Additionally, the preference may be due to the participant's prior experience and expertise with the list interface.

CONCLUSION

There were no significant differences found in performance measures on the list and tabular interfaces. The performance can be considered as influenced by their prior experiences with the search result interfaces that present data in a linear order. However, the eye movement's results imply that search result interfaces should take into account viewing patterns of its users, and thereby, design interfaces that optimize scanning strategies. The eye movement's data provided insights into the ocular behavior of users in

scanning textual data spaces. The eye movement's data supplemented with the conventional usability measures can help in diagnosing interfaces, and developing user centered designs. Future work for this research can use color coding for results, provide flexibility by allowing users to manipulate size and date categories of the results according to the context of the task.

REFERENCES

1. Broder, a. (2002). A taxonomy of web search. SIGIR forum, 36(2):3-10.
2. RealNames, 2000 http://www.realnames.com
3. Forrester report (October 2000) http://www.forrester.com
4. AllTheWeb http://www.alltheweb.com
5. Bandos J., & Resnick M. L., Improving user search with embedded Boolean search hints. In *Proc HFES 2004*. 48, 1523-1527
6. Goldberg J. & Kotval X. P. Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics. 1999.* 24, 631-645.
7. Schiessl M., Duda S., Thölke A., & Fischer R. Eye tracking and its application in usability and media research
8. Poynter Eye-tracking Research http://www.poynterextra.org/eyetrack2004/about.htm
9. Goldberg J**.,** Stimson M., Lewenstein M., Scott N., Wichansky A**.** Eye tracking in web search tasks: design implications In *Proc.* of the *Symposium on Eye tracking research & applications.* 2002. 51-58.
10. Resnick M. L., Maldonado C., Santos J., Lergier R. Modeling On-Line Search Behavior Using Alternative Output Structures. *In Proc. HFES 2001,* 1166-1170.
11. Dumais S., E. Cutrell and H. Chen. Bringing order to the web: Optimizing search by showing results in context. In *Proc. CHI'01, Human Factors in Computing Systems,* April 2001, 277-283.
12. 12. Resnick M. L., & Bandos J. Best practices in search user interface design. In *Proc*. *HFES 2002* 42*,* 627-631.
13. Granka L., Joachims T., Gay G., Eye-tracking analysis of user behavior in WWW search. In *Proc* of *Research and development in information retrieval***,** ACM Press (2004), 478 – 479.
14. Pan B., Hembrook H., Gay G., Granka L., Feusner M., Newman J. The determinants of web page viewing behavior: an eye-tracking study. In *Proc. Eye tracking research & applications symposium.* 2004. 147 - 154
15. Salogarvi, J. Kojo, I., Jaana, S., & Kaski, S. Can relevance be inferred from eye movements in information retrieval? In *Proc. of the Workshop on Self-Organizing Maps (WSOM'03), Hibikino, Kitakyushu, Japan*, 261-266
16. Klockner K., Wirschum N., Jameson A., Depth and Breadth-First Processing of Search Results Lists.
17. Jansen, B.J., Spink, A., & Saracevic. T. Real life, real users, and real needs: A study and analysis of user queries on the Web. *Information Processing and Management, 36*(2), 207-227.
18. Silverstein, C., Henzinger, M., Marais, J., Miricz, M. Analysis of a very large AltaVista query log. Technical Report, Hewlett Packard Laboratories, 1998 Number SRC-TN 1998-014, Oct.
19. Duchowski A. T. A Breadth-First Survey of Eye Tracking Applications. *BehaviorResearch Methods, Instruments, & Computers (BRMIC)*. 2002 34(4), pp.455-470.