The Effect of User Characteristics on Search Effectiveness in Information Retrieval

Azzah Al-Maskari
Directorate General of Technological Education
P.O. Box: 413, Post code 100
Tel: +96824839124
Fax: +96824813723
Email: maskeri.a@gmail.com

Mark Sanderson (corresponding author)
School of Computer Science and Information Technology
RMIT University
GPO Box 2476, Melbourne 3001
Victoria, Australia

Abstract
This paper investigates the influence of user characteristics (e.g. search experience and cognitive skills) on user effectiveness. A user study was conducted to investigate this effect, 56 participants completed searches for 56 topics using the TREC test collection. Results indicated that participants with search experience and high cognitive skills were more effective than those with less experience and slower perceptual abilities. However, all users rated themselves with the same level of satisfaction with the search results despite the fact they varied substantially in their effectiveness. Therefore, information retrieval evaluators should take these factors into consideration when investigating the impact of system effectiveness on user effectiveness.

The Effect of User Characteristics on Search Effectiveness in Information Retrieval

Abstract

This paper investigates the influence of user characteristics (e.g. search experience and cognitive skills) on user effectiveness. A user study was conducted to investigate this effect, 56 participants completed searches for 56 topics using the TREC test collection. Results indicated that participants with search experience and high cognitive skills were more effective than those with less experience and slower perceptual abilities. However, all users rated themselves with the same level of satisfaction with the search results despite the fact they varied substantially in their effectiveness. Therefore, information retrieval evaluators should take these factors into consideration when investigating the impact of system effectiveness on user effectiveness.

1. Introduction

Information retrieval is a complex activity involving interactions between users and the systems that provide the information. However, users have many factors which may influence their effective use of the information retrieval (IR) systems; for example, users experienced with information retrieval may make better and more effective use of an IR system as compared to less experienced users. Users of IR systems also employ a variety of cognitive processes when they retrieve, including learning, comprehension and speed in spotting information. All these factors contribute to user's effectiveness of the search process. It seems likely therefore that individual differences in users can affect their information retrieval effectiveness despite searching the same system and the same tasks.

There have been a few attempts to investigate the correlation between user effectiveness and system effectiveness (e.g. Hersh et al., 2000; Turpin and Hersh, 2001; Allan et al., 2005; Turpin and Scholer, 2006; Smith and Kantor, 2008); however these studies did not establish any direct correlation, finding only that improvements in IR system effectiveness (as quantified by IR effectiveness measures) do not always translate into a direct benefit for end-users (as measured by the number of relevant documents obtained and the time taken to save the first relevant document). However, it must be noted that the assessment procedures used in these studies did not take into consideration variation in personality and cognitive characteristics among users; instead, these studies used analytical techniques which assumed that all users exhibited the same or similar characteristics. However, Marchionini (1995: p.75) refuted such an assumption, asserting that systems be designed and evaluated based on a clear knowledge of the targeted user group's characteristics and their search tendencies. Therefore, any comprehensive user evaluation protocol requires not only examining how users interact with retrieval systems in the context

of a specific search task, but also a holistic view of the elements which compose the users' entire information-seeking experience. This is also described more recently in Ingwersen & Jarvelin (2005). It is not sufficient, therefore, to merely seek to establish a correlation between user effectiveness and system effectiveness; it is also important to analyze the user as a complex individual entity with a unique repertoire of cognitive skills and motivations which directly affect the execution and outcome of any search task.

Research in human computer interaction has demonstrated that primary cognitive abilities represent a powerful predictor of information-seeking success in information systems (Downing et al., 2005). The importance of examining the entire information-seeking experience of the user (e.g. knowledge, capabilities, expertise, search experience and searching behaviour) when evaluating an IR system was recommended by several researchers (e.g. Harman, 1992; Salton, 1992; Kim and Allen, 2002).

Several studies investigated the effect of user attributes on effectiveness (e.g. Marchionini, 1995, McDonald and Stevenson, 1998, Patel et al., 1998, Allen, 1992; Palmquist and Kim, 2000; Kim and Allen, 2002; Juvina and van Oostendorp, 2004; Downing et al., 2005). These studies focused on information search tasks, which shared characteristics with many electronic information environments that relied heavily on navigation (e.g., hierarchical databases, the World Wide Web, FirstSearch, etc.). For example, Allen (1992) required users to search for references for a term paper using a standard CD-ROM based index. Palmquist and Kim (2000) required users to carry out factual[1] and topical[2] search tasks on a University Web site. Kim and Allen, (2002) assigned participants two search tasks of different types: a known-item[3] search and a subject search[4] task. Juvina and van Oostendorp (2004) used web navigation tasks. Downing et al.'s (2005) study required each participant to search for information related to five tasks using FirstSearch software. These tasks included one neutral search, two business oriented tasks, and two biology oriented tasks. The main focus of a number of studies was investigating effects of task on information search behaviour. Therefore, this paper is designed to investigate the effect of user characteristics on effectiveness while searching for the less examined recall-oriented task 'to identify as many relevant documents' for a given set of topics. Furthermore, the previous studies are not recent and it is worth revisiting the impact of user characteristics on effectiveness in the context of Web search.

This study examines how individuals with different characteristics effectively complete a search task. It is hypothesized that a users' search experience and perceptual speed influences user effectiveness. This

[1] Find information on general requirements for applying for graduate study at the University.
[2] Find information helpful to the participant's future career plans. As the participant was allowed to search for information related to the career of his or her interest, the relevance of the retrieved information was judged subjectively by the participant.
[3] Find a piece of information that is known to exist in the collection being searched.
[4] Find different pieces of information that are related to the subject given and regarded as useful to the searcher.

paper is divided as follows: definition of user characteristics and related research is provided in section 2. Section 3 presents the experimental design including participants, the system, the task, and perceptual speed tests. Section 4 presents the results on the effect of user characteristics on user effectiveness and section 5 presents a discussion and the conclusion of this work.

2. Definition of user characteristics and related research

Human factors and individual differences were recognized as a key aspect for understanding user search behaviours (Egan, 1988; Nielsen, 1993). Some of these factors were closely related to user characteristics/traits such as motivation, and experience in various aspects such as computing, librarianship, and skills in searching for information. Marchionini (1995) explained that every individual has a unique set of IR skills; he described these skills, experiences and resources as a 'personal information infrastructure' (p.11). The infrastructure concept consists of three components: domain, system, and search expertise. The concept "user characteristics" – as used in this paper – refers to user's search experience and cognitive skills.

2.1. Experience

Several studies found correlations between user's search effectiveness and searching experience. Sullivan et al. (1990) found that end users (doctoral students) retrieved fewer items than experts (skilled librarians). Lancaster et al. (1994) compared the results of searches completed by librarians with those done by end users; they found that librarian searches tended to be higher quality (i.e., fewer errors and higher recall). Haynes et al. (1994) compared the search results of clinicians and experienced librarians and concluded that clinicians achieved fewer precise results. The librarians tended to conduct focused successful searches. Saito and Miwa (2001) found that experience affected information seeking behaviour on the Web; significant differences between expert and novice users were found in relation to time spent on task completion.

Lazonder et al. (2000) also studied the effects of levels of experience when searching the Web. Twenty-five pre-university students performed three tasks, each designed to observe their performance in locating Web sites and finding information contained within those Web sites. The first part of each task entailed skills in using a search engine; the second part required browsing skills. Students with higher levels of Web experience displayed superior searching, requiring less time, producing more correct responses, and engaging in fewer actions.

2.2. Users' cognitive skills

In addition to search experience discussed above, cognitive factors may potentially influence the effectiveness of the user's information retrieval process. Some studies investigated the effect of users' cognitive abilities (e.g., perceptual speed, logical reasoning, verbal comprehension, and spatial scanning) on their searching effectiveness. Allen (1992) found that cognitive speed and cognitive abilities such as logical reasoning and verbal comprehension influenced information seeking in a CD-ROM bibliographic search task. He concluded that users who scored higher on standard tests of perceptual speed performed higher quality searches than those who score lower on these tests. Palmquist and Kim (2000) studied the effects of cognitive style[5] on user effectiveness. The researchers found that while cognitive style significantly influenced the search performance of novice searchers, the influence was greatly reduced for those searchers who had substantial experience searching on-line databases. Novice searchers with low cognition tended to spend more time and visited more links for retrieving relevant information than those with higher cognition. Papaeconomou et al (2008) followed up the Palmquist and Kim research with a small study on search and learning style.

Participants with online search experience, yet low cognition, tended to visit a lower number of links for retrieving relevant information than those with no online search experience. Downing et al. (2005) also examined the effect of spatial scanning in a hierarchical database system. They concluded that individuals with low spatial scanning took longer to find the first relevant article and obtained fewer relevant articles than those with high spatial scanning. Kim and Allen (2002) also found that in web search engines, cognitive abilities (analytic cognitive styles) influenced the number of tasks completed, the number of sites visited, the number query terms issued and the number of bookmarks made.

3. Experiment methodology

Our study was designed to investigate the effect of user characteristics (e.g. search experience and cognitive skills) on their effectiveness. A large heterogeneous group of users with different characteristics were recruited and were asked to search for a given set of topics; then their effectiveness was compared against their characteristics.

The TREC[6] test collections were used as a vehicle for providing a better understanding of retrieval evaluation since the 1992 (Voorhees, 2000). There were little research examining the relationship between user characteristics and their search effectiveness while searching TREC test collections. This is an important omission as a number of studies were conducted on TREC data examining the correlation

[5] Defined as exhibited preferred modes or strategies detectable as distinctive or characteristic methods of performing.
[6] Text REtrieval Conference http://trec.nist.gov/ [site accessed: 23/01/08].

between test collection based effectiveness scores and user performance (e.g. Hersh et al., 2000; Turpin and Hersh, 2001; Allan et al., 2005; Turpin and Scholer), but none of these studies examined the impact of user characteristics in these correlations. To test this effect, an interactive experiment was conducted by recruiting 56 users searching for 56 TREC topics; the topics are listed in the appendix in Section 7.

Users completed searching for 56 topics by utilizing an experimental test bed (Query Performance Analyzer)[7]. This system allowed access, through a single consistent interface, to three well known experimental retrieval systems (InQuery, Lemur and Terrier). The 56 topics were randomly selected from TREC topics. For more details description of the system, see Al-Maskari et al.(2008).

3.1. Participants

Fifty-six participants were recruited for this experiment and were required to save as many relevant documents as possible for a set of 56 TREC topics. (Note, have the same number of topics as users was a coincidence.) Every user completed searches for 8 topics and every eight users completed searches on the same set of topics. Users were given 7 minutes for each topic, and topics were randomly distributed amongst users (to reduce the effects of topic order on results). Upon completing a search for each task, users rated their satisfaction with the results in a 4-point scale: *very satisfied,partially satisfied, fairly dissatisfied*, and *very dissatisfied*. Users were presented with the description and narrative fields of TREC topics as information needs to be satisfied. They were free to issue multiple queries for each topic within the 7 minutes. The narrative field served as guidance on assessing document relevance using a ternary relevance scheme: highly relevant[8], partially relevant[9] or not relevant. Prior to commencing the experiment series, every user underwent a training session consisting of two training tasks to ensure familiarity with the system and experimental protocols.

The participants were required to fill out self-evaluation questionnaires including demographic data (such as gender, age, academic majors, degrees earned and current academic program) and rated their general computer literacy and experience using search engines. Of the 56, 26 were male and 30 were female. All participants were, at the time of participation, students at the University of Sheffield in different faculty with 20 being undergraduates, 29 reading for MSc degrees and 7 PhD students. As shown in Table 1 the students came from various academic backgrounds. Figure 1 shows the distribution of age. The participants' computer experience, as measured by their responses to self-assessment questionnaire, is presented in Figure 2[10]. Only three subjects rated themselves as below average computer literacy skills.

[7] http://www.info.uta.fi/julkaisut/pdf/qparn1.pdf [site accessed: 22/02/08].
[8] The document directly addresses the core issue of the topic.
[9] The document only points to the topic: it does a not discus the themes of the topic thoroughly.
[10] Subjects indicated their level of expertise with computers and internet usage on a seven point Likert scale.

Topic	#	Topic	#	Topic	#
Information studies	9	Health related studies	5	Sociology	3
Misc social sciences	9	Comp Sci/Maths	5	Linguistics	3
Engineering	8	History	4		
Medicine	7	Business	3		

Table 1: Breakdown of participants' areas of study

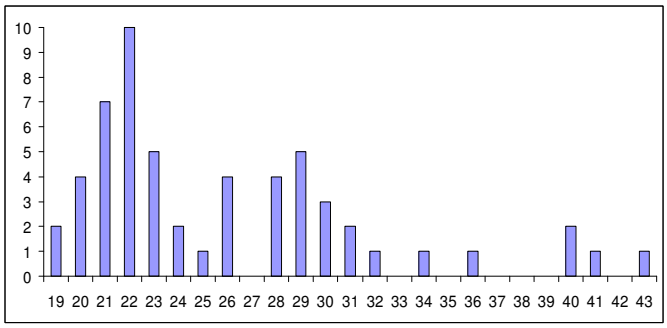

Figure 1: Break down of subjects' age.

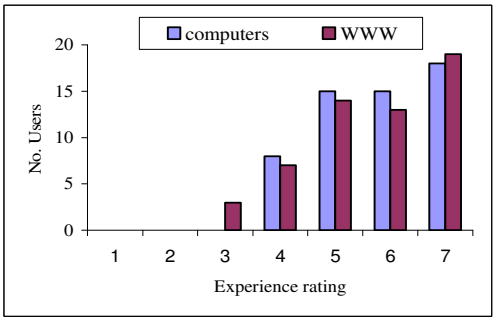

Figure 2: Subjects' experience (1=very little; 7=high).

To provide a more objective assessment of their experience with using online search engines, subjects were asked to indicate the number of years that they had been using online search. These results are displayed in Figure 3. Overall, subjects had a mean of 6.7 years of such online search experience, with one individual claiming 12 years of search engine usage; conversely, there was one recruit whose experience of interacting with online IR systems was limited to just one year. On average, 18% of the subjects had less than five years experience, 45% had between five to seven years experience while 36% boasted between eight to ten years experience. Subjects were also asked to indicate the number of hours per day they spend conducting online searches, the statistics for which are displayed in Figure 4. Overall subjects had a mean of 3.7 hours search per day ; 63% of them claimed to spend two to three hours daily using online search engines, with one individual averaging 14 hours per day.

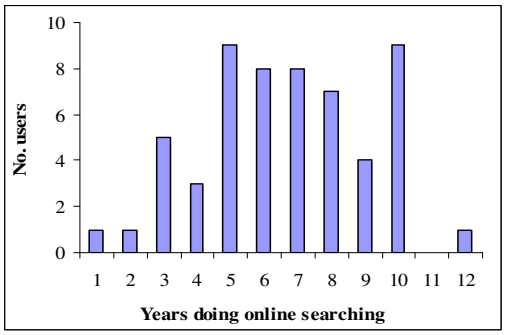

Figure 3: User experience in terms of number of years doing online search

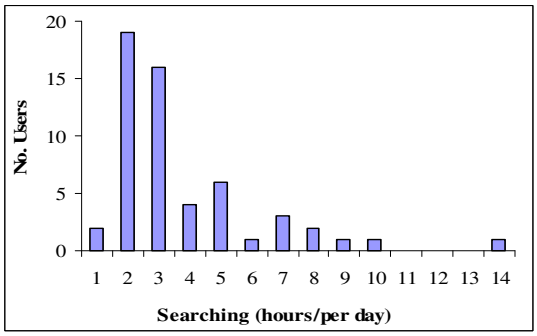

Figure 4: User experience in terms of search per day (hours).

3.2. Assessing Users' Perceptual Speed

All participants were initially assessed to determine their perceptual speed. Participants completed three tests of perceptual speed from the suite of evaluation exercises known as the Factor-Referenced Cognitive Tests (FRCT), produced by the Educational Testing Service[11] (Ekstrom et al., 1976). For the experiment described here, the cognitive ability of interest is known as 'perceptual speed', which is deemed to be one of the principal factors influencing informational retrieval behaviour (Allen, 1992; Kim and Allen, 2002; Palmquist and Kim, 2000). The specific tests used to assess the users' perceptual speed were

(i) the "Finding A" Test, in which participants identified words that had the letter "a"; the test consisted of four pages (each page contained 205 words to be inspected) to be completed within 2 minutes

(ii) the "Number Comparison" Test, in which participants inspected pairs of multi-digit numbers and indicated whether the two numbers in each pair were identical or otherwise; the test consisted of one page (46 pairs of numbers) to be completed within 1½ minutes

(iii) the "Identical Pictures" Test, in which participants examined a sequence of five numbered geometrical figures or pictures and indicated whether they were the same as a reference sequence; the test consisted of two pages (each page with 24 objects that required a match) to be completed within 1½ minutes

Participants who finished a test before the time limit had expired were permitted to proceed to the next test. The maximum overall time allowed to complete the tests was 5 minutes. The order of test presentation was randomized to minimize any possible learning effects. Written instructions and practice exercises were presented immediately before the participants completed each test.

3.3. Variables Tested in this experiment

As mentioned earlier, this study aimed at identifying the effect of user characteristics on user effectiveness. Therefore, user effectiveness in this experiment was measured by four variables:

- UserDocs – the number of relevant documents the users identified.
- TimeFRD – the time taken by the user to locate the first relevant document.
- Satisfaction – A user self-assessment of their satisfaction with the search for a topic
- Familiarity – A user self-assessment of their familiarity with the search topic.

[11] http://www.ets.org

4. Results

The results section consists of four parts: the effect of users' search experience on their effectiveness; the effect of perceptual speed on user effectiveness; users' satisfaction with the search results; and users' familiarity with the search topics.

4.1. Effects of Search Experience on User Effectiveness

As described in section 3.1 all participants recruited were required to fill out self-evaluation questionnaires, rating their general computer literacy and experience using search engines. Table 2 presents a Pearson's correlation analysis between search effectiveness and experience. Of the various attributes of search experience considered (e.g. years of searching, number of hours spent online using search engines, frequency of using computer, and frequency of searching online), the only attribute showing any correlation with user effectiveness was user's search experience though the correlation was low. Table 3 provides statistical data concerning the search experiment of recruited users and their search effectiveness.

	UserDocs	TimeFRD
Years-searching	0.11*	-0.049
Hours-searching (daily)	0.178	0.155
Frequency computer usage	0.078	0.001
Frequency online searching	-0.046	-0.043

Table 2: Pearson's correlation between user effectiveness and search experience (*p<0.05.).

	Range	Min	Max	Mean	Std. Dev
Years-searching	11.00	1.00	12.00	6.47	2.60
Hours-searching- (daily)	13.00	1.00	14.00	3.84	2.50
Computer-frequency	3.00	4.00	7.00	6.29	0.84
Online search frequency	4.00	3.00	7.00	5.73	1.01
UserDocs	7.38	2.13	9.50	4.56	1.73
TimeFRD	3.36	0.77	4.13	1.95	0.74

Table 3: Descriptive statistics of users search experience and their effectiveness.

Since years of search experience is the only attribute which correlates at any level with user effectiveness, an investigation was conducted to confirm whether users with more years of online search experience were consistently finding a higher number of relevant documents per search task than those with less experience. A split analysis was performed on the "years-searching" attribute: users with $\leq$7 years were classified as less experienced (N=35); those with >7 years were considered experienced users (N=21). Results (Figure 5 and Figure 6) reveal that, on average, less experienced users scored less in *UserDocs* and more in *TimeFRD* than their more experienced counterparts. This was again confirmed by Mann-

Whitney test[12] of the same data which indicated a statistically significant difference in *UserDocs* scores between experienced users, 5.2 docs, and inexperienced users, 4.2 docs, (see Figure 5, $U = 20526, N_1 = 280, N_2 = 168, p = 0.023$, two-tailed). However, experienced users, averaging 1.84 minutes, did not significantly outperform their less experienced counterparts, who averaged 2.11 minutes (see Figure 6, $U = 12695, N_1 = 168, N_2 = 167, p = 0.133$, two-tailed).

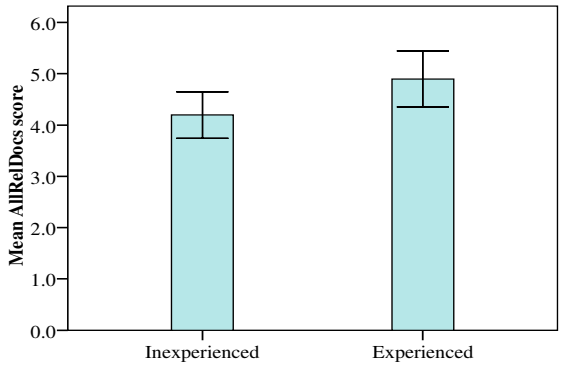

Figure 5: UserDocs score compared with users' search experience.

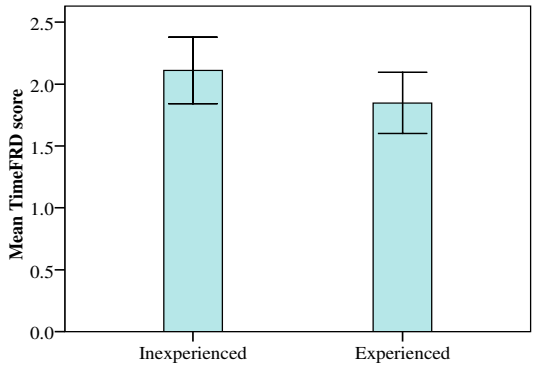

Figure 6: TimeFRD score compared with users' search experience.

4.2. The Effects of Perceptual Speed on User Effectiveness

Table 4 illustrates the Pearson's correlation between the three perceptual speed tests and user search effectiveness (*UserDocs* and *TimeFRD*). This table shows that participant performance in the Identical Picture Test exhibited a stronger Pearson's correlation than the other two tests respectively. The data presented in Table 4 consists of the search effectiveness for each user; the user effectiveness averaged over 8 search topics.

Perceptual speed tests	UserDocs	TimeFRD
'A' letter test	0.062	-0.174
No. comparison test	0.235	-0.317**
Identical pictures test	0.310*	-0.408**
Overall PS (sum of the above three tests)	0.258	-0.383**

Table 4: Pearson's correlation between users' effectiveness and their perceptual speed; **p<0.01; *p<0.05

Given the weaker correlation of the other two tests vis-à-vis user effectiveness, it was decided to formulate a measure called as "Overall Perceptual Speed (OPS)" incorporating all three components. To ensure that all three tests were actually assessing the same ability, *reliability analysis* (de Vaus, 2002) was

[12] A non-parametric significance test used for comparing two independent variables.

employed using Cronbach's alpha[13]. The alpha value for the A's Test was $\alpha = 0.79$, $\alpha = 0.73$ for the Number Comparison test and $\alpha = 0.83$ for the Identical Picture Test, giving an overall Cronbach's reliability rating of $\alpha = 0.87$, above the 0.7 threshold value (De Vaus, 2002: p.20).

As with user experience, participants were assigned to a low or high group, based on a median split of perceptual speed scores. Table 5 presents a statistical comparison of the users' effectiveness in relation to their perceptual speed from Table 4.

	UserDocs		TimeFRD	
	High PS	Low PS	High PS	Low PS
Average	4.73	4.39	1.71	2.19
St.dev	1.68	1.73	0.68	0.70

Table 5: Overall perceptual speed and search effectiveness (from Table 4)

A Mann-Whitney test was conducted to examine whether user effectiveness varied significantly depending on a users' perceptual speed. Although Pearson's correlation analysis does indicate a significant variation, this is not evident when Mann-Whitney test is applied for the *UserDocs* score: ($U = 3412, N = 28, N = 28, p = 0.408,$ two tailed). Perceptual speed seems to only influence the *TimeFRD* score ($U = 216, N = 28, N = 28, p = 0.004,$ two tailed); indicating that users with lower perceptual speed took longer to obtain the first relevant document than those with higher perceptual speed. Figure 7 illustrates the user effectiveness in relation to (*TimeFRD*) their perceptual speed scores.

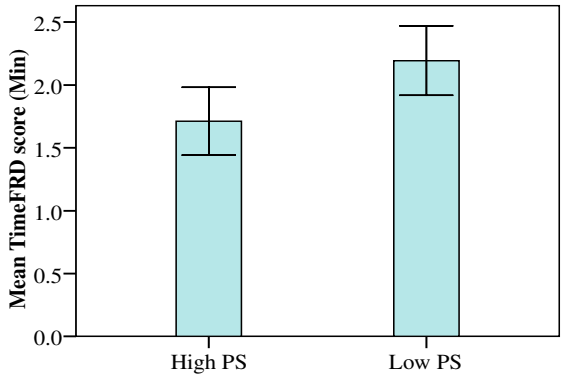

Figure 7: Effect of user perceptual speed on TimeFRD score as demonstrated via: (a) split of users with high or low perceptual speed.

4.3. Users' satisfaction with the search results

This section examines the relationship between user characteristics and user satisfaction, the effect of perceptual speed and online search experience on user satisfaction. The results in Figure 8 showed that

[13] A measure of how well a set of items (or variables) measures a single dimensional latent construct.

users with high perceptual speed were equally satisfied as users with low perceptual speed ($U = 24422, \mathrm{N}_1 = 248, \mathrm{N}_2 = 200,\ p = 0.772$). A similar situation existed for experienced and inexperienced user satisfaction ($U = 22618, \mathrm{N}_1 = 280, \mathrm{N}_2 = 168,\ p = 0.477$) in Figure 9.

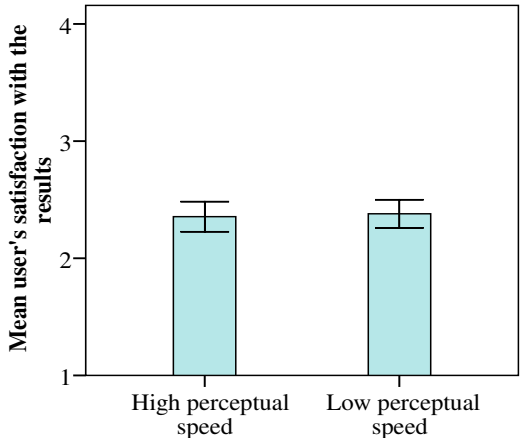

Figure 8: Users' satisfaction in relation to perceptual speed.

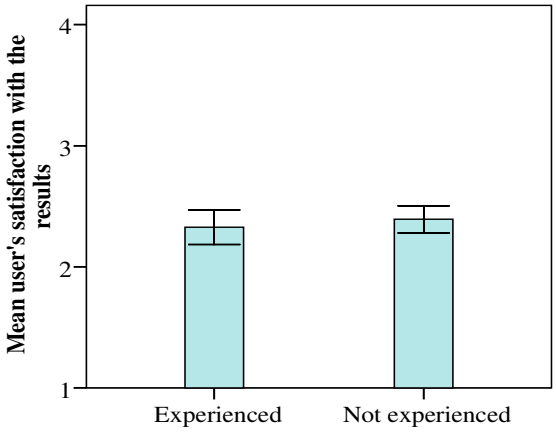

Figure 9: Users' satisfaction in relation to online search experience.

4.4. Users' familiarity with the search topics

After completing the search for each topic, users rated their familiarity using a 4-point Likert scale: (i) familiar, (ii) fairly familiar, (iii) fairly new, and (iv) very new. This post-search questionnaire was also used in Kelly and Cool (2002). A Spearman's correlation between users' perceptual speed and familiarity showed that there was no significant correlation between the two (r=0.-0.038, p=0.57). Similarly, there was no correlation between years of searching experience and familiarity (r= 0.199, p=0.14).

5. Discussions and Conclusions - Factors Validation

The study reported in this paper revisited the effect of user searching experience on user effectiveness. In this paper it was demonstrated that users with more than seven years of online searching experience obtained significantly more relevant documents (5.2 relevant documents) than users with less experience (4.2 relevant documents). These results are consistent with previous research which demonstrated that domain expertise and searching experience enhanced user effectiveness (Marchionini, 1995; McDonald and Stevenson, 1998; Patel et al., 1998). However, in our study, experienced and non-experienced users took nearly the same amount of time to obtain the first relevant document; they only varied in the number of relevant documents obtained.

Furthermore, this study also demonstrated a significant correlation between users' effectiveness and their perceptual speed: users who scored above the median of the perceptual speed tests took significantly less time to obtain the first relevant document (1.71 minutes) than users who scored below the median of these tests (2.19 minutes). Therefore, users with high cognitive skills can spot information with speed and accuracy. Using different search tasks and different IR system, the results from this experiment are consist with the findings reported by previous research (e.g. Allen, 1992; Palmquist and Kim, 2000; Kim and Allen, 2002; Juvina and van Oostendorp, 2004; Downing et al., 2005). However, in our study there was no significant difference in the total number of relevant documents obtained by users with different perceptual speed skills.

Furthermore, as stated in section 3.2, users' perceptual speed was assessed based on three tests: the "Finding A's"; the "Number Comparison" Test, and the "Identical Pictures" test. From this experiment, we found that the identical picture test impacted more on user effectiveness than the number comparison test whereas Allen (1992) found the opposite. Perhaps this disparity is due to the different tasks involved and different systems used in this study and in Allen's study (CD-ROM index and searching for articles related to the 'prescription for reduction of aggression').

However, it is surprising to find that users with high and low perceptual speed rated themselves with the same level of satisfaction despite the fact that users with high perceptual speed obtained the first relevant document faster than users with low perceptual speed. Furthermore, experienced and inexperienced users rated themselves with the same level of satisfaction, although they varied significantly in the number of relevant documents obtained. The reason for the lack of difference in user satisfaction between the two groups was most likely due to the subjective nature of user satisfaction and the differences in users' opinion and perception of the search results. Furthermore, user satisfaction can be easily influenced by several factors, (e.g., system effectiveness, user effectiveness, and user characteristics and expectations) as was addressed by Al-Maskari et al (2010).

In conclusion, the findings from this study reinforced the fact that users of IR systems employ a variety of cognitive processes when retrieving information, including learning, comprehension and speed in spotting information. All these factors contributed to user's effectiveness of the search process. Therefore, it was not sufficient to merely seek to establish a correlation between user effectiveness and system effectiveness; it was also important to analyze the user as a complex individual entity with a unique repertoire of cognitive skills and motivations which directly affected the execution and outcome of any search task.

In light of the findings above, it was reasonable to argue that certain user characteristics can be used to predict search effectiveness or to explain differences in effectiveness between user groups. These results

have implication in the evaluation of interactive IR systems. These results also have important practical implications concerning the design of IR systems which must take into consideration the influence of a user's cognitive ability in determining search behaviour. Such investigations of cognitive abilities can provide fruitful input to the design process. These results may lead to the development of a variety of search system features to augment cognitive abilities. A system embedded with these features can contribute to the effective use of information retrieval systems by people with different levels of cognitive abilities and aptitudes.

This study obtained users' familiarity with the search topics after completing the search tasks. Future research will ask users to rate their familiarity with the search topics before and after the search. This approach would allow us to measure how much knowledge the users gained upon reading the information retrieved by the IR system and therefore how much new additional information was provided to them.

6. References

Al-Maskari, A., Sanderson, M. & Clough, P. (2008)."The Good and the Bad System: Does the Test Collection Predict Users' Effectiveness". In: *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pp.59-66 Singapore: ACM.

Al-Maskari, A., Sanderson, M. (2010). A review of factors influencing user satisfaction in information retrieval. *Journal of the American Society for Information Science and Technology*. **61**(5) , 859-868

Allan, J., Carterette, B., & Lewis, J. (2005). When will information retrieval be "good enough"? *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 433-440). ACM New York, NY, USA.

Allen, B. (1992)."Cognitive differences in end user searching of a CD-ROM index". In: *Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval*, pp.298-309. Copenhagen, Denmark: ACM.

de Vaus, D. (2002) Analyzing Social Science Data, London, Sage Publications Inc.

Downing, R. E., Moore, J. L. & Brown, S. W. (2005) The effects and interaction of spatial visualization and domain expertise on information seeking. *Computers in human behavior,* 21 (2), 195-209.

Egan, D. E. (1988) Individual differences in human-computer interaction. IN HELANDER, M. G., LANDAUER, T. K. & PRABHU, P. V. (Eds.) *Handbook of human-computer interaction* Amsterdam, Holland, Elsevier.

Ekstrom, R. B., French, J. W., Harman, H. H. & Dermen, D. (1976) Manual for kit of factor-referenced cognitive tests, Princeton, NJ, Educational Testing Service.

Harman, D. (1992) Evaluation Issues in Information Retrieval. *Information Processing and Management,* 28 (4), 439-40.

Haynes, R. B., Walker, C. J., McKibbon, K. A., Johnston, M. E. & Willan, A. R. (1994) Performances of 27 MEDLINE systems tested by searches with clinical questions. *Journal of the American Medical Informatics Association,* 1 (3), 285-295.

Hersh, W., Crabtree, M., Hickam, D. & Friedman, C. (2000) Factors associated with successful answering of clinical questions using an information retrieval system. *Bull Med Library Association,* 88 (4), 323-331.

Ingwersen, P., & Järvelin, K. (2005). *The Turn: Integration of Information Seeking and Retrieval in Context*. Springer.

Juvina, I. & van Oostendorp, H. (2004)."Individual Differences and Behavioral Aspects Involved in Modeling Web Navigation". In: *Proceedings in the 8th ERCIM Workshop "User Interfaces For All"* pp.77-95. Vienna, Austria: Springer

Kelly, D., & Cool, C. (2002). The effects of topic familiarity on information search behavior. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries* (pp. 74-75). ACM.

Kim, K.-S. & Allen, B. (2002) Cognitive and task influences on Web searching behavior. *Journal of the American Society for Information Science and Technology,* 53 (2), 109-119.

Lancaster, F. W., Cheryl, E., Mary, J. Z., Laura, M. & Yuen-Man, L. (1994) Searching Databases on CD-ROM: Comparison of the Results of End-User Searching with Results from Two Modes of Searching by Skilled Intermediaries. *RQ,* 33 (3), 370-386.

Lazonder, A. W., Biemans, H. J. A., & Wopereis, I. G. J. H. (2000). Differences between novice and experienced users in searching information on the World Wide Web. Journal of the American Society for Information Science, 51(6), 576-581.

Marchionini, G. (1995) Information Seeking in Electronic Environments Cambridge University Press.

McDonald, S. & Stevenson, R. J. (1998) Navigation in hyperspace: An evaluation of the effects of navigational tools and subject matter expertise on browsing and information retrieval in hypertext. *Interacting with Computers,* 10 (2), 129-142.

Nielsen, J. (1993) Usability Engineering, San Francisco, Morgan Kaufmann.

Palmquist, R. A. & Kim, K. S. (2000) Cognitive style and on-line database search experience as predictors of Web search performance. *Journal of the American Society for Information Science,* 51 (6), 558-566.

Papaeconomou, C., Zijlema, A. F., & Ingwersen, P. (2008). Searchers' relevance judgments and criteria in evaluating web pages in a learning style perspective. In *Proceedings of the second international symposium on Information interaction in context* (pp. 123-132). ACM.

Patel, S. C., Drury, C. C. & Shalin, V. L. (1998) Effectiveness of expert semantic knowledge as a navigational aid within hypertext. *Behavior and Information Technology,* 17 (6), 313-324.

Saito, H. & Miwa, K. (2001)."A Cognitive Study of Information Seeking Processes in the WWW: The Effects of Searcher's Knowledge and Experience". In: *Proceedings of the Second International Conference on Web Information Systems Engineering (WISE'01) Volume 1*, pp.321-327. Washington, DC, USA: IEEE Computer Society.

Salton, G. (1992) The State of retrieval system evaluation. *Information Processing and Management,* 28 (4), 441-449.

Smith, C. L. & Kantor, P. B. (2008)."User adaptation: good results from poor systems". In: *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pp.147-154 Singapore: ACM.

Sullivan, M. V., Borgman, C. L. & Wippern, D. (1990) End-users, mediated searches, and front-end assistance programs on Dialog: A comparison of learning, performance, and satisfaction. *Journal of the American Society for Information Science,* 41 (1), 27-42.

Turpin, A., & Hersh, W. (2001). Why batch and user evaluations do not give the same results. *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 225-231). ACM New York, NY, USA.

Turpin, A., & Scholer, F. (2006). User performance versus precision measures for simple search tasks. *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 11-18). ACM New York, NY, USA.

Voorhees, E. M. (2001) The philosophy of information retrieval evaluation. *Evaluation of Cross-Language Information Retrieval Systems, Second Workshop of the Cross-Language Evaluation Forum, CLEF*, 9–26.

7. Appendix

Here the list of TREC topics used in the experiment is provided grouped into three broad categories.

Health/Medicine

1. Evidence that radio waves from radio towers or car phones affect brain cancer occurrence.
2. Is the disease of Poliomyelitis (polio) under control in the world?
3. What research has been done on viral hepatitis and what progress has been made in its treatment?
4. What drugs are being used in the treatment of Alzheimer's Disease and how successful are they?
5. Is the fear of open or public places (Agoraphobia) a widespread disorder or relatively unknown
6. What are the advantages and/or disadvantages of tooth implants?
7. What adverse effects have people experienced while taking aspirin repeatedly?
8. Is it hazardous to the health of individuals to work with computer terminals on a daily basis?
9. Identify documents discussing the use of estrogen by postmenopausal women in Britain
10. Identify documents discussing cases where rabies have been confirmed and what, if anything, is being done about it
11. Identify documents that discuss in vitro fertilization.
12. What role does blood-alcohol level play in automobile accident fatalities?
13. What are the benefits, if any, of drug legalization?

14. What forms of alternative medicine are being used in the treatment of illnesses or diseases and how successful are they?
15. Identify documents that discuss sick building syndrome or building-related illnesses.
16. What is happening in the field of behavioral genetics, the study of the relative influence of genetic and environmental factors on an individual's behavior or personality?
17. How widespread is carbon monoxide poisoning on a global scale
18. Find documents that discuss the damage ultraviolet (UV) light from the sun can do to eyes.
19. Do any countries other than the U.S. and China have a declining birth rate?
20. Identify drugs used in the treatment of mental illness.
21. Identify documents that discuss medical treatment of obesity.
22. Identify documents that discuss the renewed popularity of cigar smoking.
23. Identify documents that discuss mercy killings.

Science

1. Commercial harvesting of marine vegetation such as algae, seaweed and kelp for food and drug purposes.
2. Identify systematic explorations and scientific investigations of Antarctica, current or planned.
3. Identify documents discussing the development and application of spaceborne ocean remote sensing.
4. What effects have been attributed to El Nino?
5. Identify documents that discuss the concerns of the United States regarding the export of encryption equipment.
6. Isolate instances of fraud or embezzlement in the international art trade.
7. Identify documents that discuss the building of a space station with the intent of colonizing the moon.
8. Identify documents that discuss effective and safe ways to permanently handle long-lived radioactive wastes.
9. Identify documents that discuss the current status of hybrid automobile engines, (i.e., cars fueled by something other than gasoline only).
10. Identify documents that discuss the use of organic fertilizers (composted sludge, ash, vegetable waste, microorganisms, etc.) as soil enhancers
11. What are the applications of robotics in the world today?
12. Identify documents that discuss the activities or equipment of oceanographic vessels
13. What unexpected or unexplained cosmic events or celestial phenomena, such as radiation and supernova outbursts or new comets, have been detected?
14. What is the impact of poaching on the world's various wildlife preserves
15. What tropical storms (hurricanes and typhoons) have caused significant property damage and loss of life?
16. What are the latest developments in robotic technology?
17. Identify instances in which weather was a main or contributing factor in the loss of a ship at sea.

Policy

1. Identify incidents of human smuggling.
2. Identify documents discussing international boundary disputes relevant to the 200-mile special economic zones or 12-mile territorial waters subsequent to the passing of the "International Convention on the Law of the Sea".
3. Identify documents that discuss the growth of Native American casino gambling.
4. Identify documents that discuss opposition to the introduction of the euro, the European currency.
5. What measures are being taken by local South American authorities to preserve the Amazon tropical rain forest?
6. How often were the peace talks in Ireland delayed or disrupted as a result of acts of violence?
7. Who is involved in the Schengen agreement to eliminate border controls in Western Europe and what do they hope to accomplish?

8. How much sugar does Cuba export and which countries import it?
9. What is known about drug trafficking in the "Golden Triangle", the area where Burma, Thailand and Laos meet?
10. What is the status of The Three Gorges Project?
11. How is the disposal of industrial waste being accomplished by industrial management throughout the world?
12. In what ways have quilts been used to generate income?
13. What countries are experiencing an increase in tourism?
14. What steps are being taken by governments or corporations to eliminate abuse of child labor?
15. Find accounts of selfless heroic acts by individuals or small groups for the benefit of others or a cause.
16. What other countries besides the United States are considering or have approved women as clergy persons?