

Using a Task-Based Approach in Evaluating the Usability of BoBIs in an E-book Environment

Abstract. This paper reports on a usability evaluation of BoBIs (Back-of-the-book Indexes) as searching and browsing tools in an e-book environment. This study employed a task-based approach and within-subject design. The retrieval performance of a BoBI was compared with a ToC and Full-Text Search tool in terms of their respective effectiveness and efficiency for finding information in e-books. The results demonstrated that a BoBI was significantly more efficient (faster) and useful compared to a ToC or Full-Text Search tool for finding information in an e-book environment.

Keywords: E-book, Task-Based, BoBI, Book Index, Usability Evaluation

1 Introduction

Typically, browsing a book or searching for specific content is accomplished via a ToC or a BoBI. In a digital or e-book environment these methods may be enhanced by hyperlink features and Full-Text Search tools. For instance, the California Digital Library Ebook Task Force [1] suggested that advanced search facilities (such as Boolean, truncation, proximity, etc.) should be incorporated in e-books. In general, a ToC provides information on the organisational structure of a book which may be skimmed to obtain some general ideas as to the book's content. A BoBI, on the other hand, provides more specific information on relevant sections of text by pointing to key concepts discussed in the book. However, although a ToC helps readers to browse through a book it relies on the reader's ability to interpret the section headings contained in the ToC. Some advantages of a BoBI compared to ToC are that it organises the information in the book into an alphabetical structure, groups together information that is scattered through the book, distinguishes important topics from random occurrences of information, and provides cross-references to indicate preferred and related terms.

In an e-book environment, we still need a BoBI even though we may already have a Full-Text Search tool. This is because a BoBI directly identifies significant topics in the book unlike a Full-Text Search tool which matches word strings specified by users. Previous research has found that users would like BoBIs to be incorporated in e-books [2-6] and this has been supported by the experiences of the Bureau of National Affairs in the United State. When the Bureau transferred its publications from paper to CD-ROM they provided Full-Text Search tools for the electronic version of their publications but excluded BoBIs [7]. However users demanded BoBIs because they were very familiar with the tool and already knew how to use it, whereas a Full-Text-Search tool required users to have a certain level of skill

and experience in order to be able to use features such as Boolean operators (AND, OR, NOT) effectively. Some other challenges raised by Full-Text Search tools are that they require users to specify search terms that match the terminology used in the text while allowing for variations in terms of spelling, hyphenation and synonyms; they retrieve and, usually, rank all instances of the occurrence of sought terms which the reader must evaluate in terms of the relevance of retrieved sections of text by browsing a list of titles.

The retrieval performance of BoBIs has been traditionally measured by researchers in terms of recall (the index finding ability) and precision (how well the index entries matched the text) [8]. With the emergence of the hypertext concept, and e-books that incorporated hyperlinks in BoBIs and ToCs, the issue of hypertext and e-book usability was introduced. Some studies such as [9-15] involved analysis of retrieval performance and user preferences between e-books or e-documents (with hypertext features) and printed books. However, the results of these studies are not directly comparable because of differences in the design of the usability tests (e.g. within subject-design or between subject design), the type of search tasks (e.g. fact finding or inference), the materials used in the evaluation (e.g. manuals, textbooks, or encyclopaedias), subject fields (e.g. chemistry or computer-based subjects), and participants (e.g. novice or computer experts) as well as interface and format issues (e.g. web or pdf versions). Therefore the study reported here may be considered significant based on the following factors:

- i. Testing was conducted using a within-subject design in which each participant was tested using each search tool (BoBI, ToC and Full-Text Search). Therefore each participant had experience of using every search tool. As a result they could provide more accurate responses on preferences and satisfaction through the interaction with the search tools. Other studies [12-14] only studied one of the search tools or let users choose the one with which they were most familiar.
- ii. The choice of e-books and search tools was randomly selected for each participant. This method was used to minimise the likelihood of users becoming familiar with the e-book content.
- iii. The study tested three different e-books from the information retrieval field. Previous studies normally used only one type of e-book for the evaluation and occasionally one book in several formats (such as [9, 12-14]).
- iv. The study involved 45 participants (students in a UK university) which is a relatively large number of participants compared to most previous studies (e.g. [13], [14] and [16]).

2 Objectives

This study was conducted with the following two main purposes:

- (i) To evaluate whether a BoBI is more effective, efficient and useful compared to a ToC and Full-Text Search tool for finding information in an e-book environment.

- (ii) To measure users' attitudes with respect to a BoBI, ToC and Full-Text Search tool for finding information in an e-book environment in terms of their preferences, levels of satisfaction and ease of use.

There were six central hypotheses that ran through the usability evaluation:

H1.1: A BoBI is more efficient compared to a ToC for finding information in an e-book environment.

H1.2: A BoBI is more efficient compared to a Full-Text Search tool for finding information in an e-book environment.

H2.1: A BoBI is more effective compared to a ToC for finding information in an e-book environment.

H2.2: A BoBI is more effective compared to a Full-Text Search tool for finding information in an e-book environment.

H3.1: A BoBI is more useful compared to a ToC for finding information in an e-book environment.

H3.2: A BoBI is more useful compared to a Full-Text Search tool for finding information in an e-book environment.

3 Research Methodology

This evaluation was carried out with subject-specific users. This was because it was assumed that the target population must have some knowledge of the subject field covered by the test collections and that they would also have reasonable and similar levels of the computer skills necessary to be able to perform a search task. The target population was MSc and research students in the Department of Computer and Information Sciences, at the University of Strathclyde and involved a total of 45 respondents with 25 of them are male and 20 are female.

This study employed a task-based and within-subject approach its evaluation design as elaborated below:

3.1 Task-Based Approach

Vakkari [17] highlighted that to characterise a search task in task-based information searching the following factors should also be taken into consideration: (i) The selection of search terms and operators in query formulation, such as the use of narrower terms, synonyms, and the use of Boolean operators such as AND, and OR; (ii) The search tactics employed, such as browsing, initial general search, and final specific search; (iii) The use of search support tools, such as query expansion for refining the query; and (iv) Relevance and utility judgements of information found, such as degree of relevance, usefulness, and precision.

In addition, typical task that users would normally undertake when interacting with e-book are browsing, searching, analysing relevant contents, and so forth, depending on the purpose for consulting the book in the first place. This depends on factors such as: (i) The types of information users search for; for example, searching for individual facts, and searching for textual or non-textual information; and (ii) The selection of search terms, such as whether users are cued by the questions that contain words in the text or headings.

The task-based approach that was utilised in this study was based on the following characteristics:

- i. The types of information searched for. This involved searching for two types of information: specific facts and relevant content. Therefore two types of search task were involved: (a) Factual task. This was a straightforward question to find a specific piece of information in the e-book (e.g. “What are the definitions of the terms precision and recall as provided in this book?”); and (b) Analytical task. This was to identify relevant e-book content which would satisfy a query about an information need involving greater breadth. For example, “You are writing an essay about some methods and implementation of automatic classification. Which section(s) of this book discuss this topic? (Give the page numbers).”
- ii. The selection of search terms in the query formulation. This study hypothesised that the appearance in the BoBI and ToC of a term in the target information probably affects users’ search performance. Therefore three types of query were formulated: (a) A term in the target information appeared only in the BoBI; (b) A term in the target information appeared only in the ToC; and (c) A term in the target information appeared in both the BoBI and the ToC.
- iii. The use of search support tools. In this case three types of search tools were used (the BoBI, the ToC and Full-Text Search) with the intent of evaluating whether one was superior to any of the others.
- iv. Relevance judgements of information found. The relevance judgement of information is influenced by users’ perceptions of past experience, their present situation, their knowledge and their search goals. In this study, relevance was constrained by the above factors and also by a predetermined correct answers set constructed by an expert in the chosen subject fields

3.2 Usability Evaluation Design

This study has employed a within-subjects approach in its usability evaluation design. The within-subject approach is also known as repeated measure where the same participants perform under all the possible combinations of conditions (in this study conditions are defined as types of search tools; i.e. BoBI, ToC and Full-Text Search). As subjects are tested under every condition, the problem of individual differences can be eliminated in which each participant acts as his/her own control. In addition, experience of using all the conditions could result in the respondents making a better judgement on each of the search tools. As repeated measures are always associated with an order effect (one task may be affected by the experience of having performed another task), or practice effect (participants’ performance on the later task may have improved since the first tasks) a procedure of counterbalancing was taken to reduce these effects. In the counterbalancing the order of the condition in within-subjects designs was varied from participant to participant in order to balance out the effects across the conditions [18]. Some advantages of this approach are that it requires fewer participants and is hence less costly, it is suitable for evaluating a system where learning is involved, and it has less chance of effects from variation between participants. This approach was employed to minimise the number of participants that

should be involved and to allow participants to interact with every search tool so that they could provide more accurate responses.

4 Results and Discussion

The data was analysed using the SPSS program. The performances of the three search tools were measured using two variables: (i) Speed (in minutes) of finding information in e-books and (ii) Count of success in finding information accurately in e-books. The data were analysed using analysis of variance (ANOVA) to test the statistical significance of the differences among the mean scores of two or more groups (in this study they were BoBI, ToC and Full-Text Search). Alternatively a Friedman Test was used to replace the ANOVA test for ordinal types of data (i.e. ranking of search tools usefulness).

4.1 Is a BoBI More Efficient Compared to a ToC and Full-Text Search Tool for Finding Information in an E-book Environment?

Fig. 1 below shows the average time (in minutes) of each search tool based on two types of search tasks. The chart shows that the BoBI outperformed the ToC and Full-Text Search for both factual and analytical tasks with an average of 3.04 and 2.44 minutes. The ToC was second best for answering factual tasks correctly (3.59 minutes) and third best for analytical tasks (4.21 minutes). The Full-Text Search on the other hand was worst for factual tasks (4.34 minutes) and second best (3.18 minutes) for analytical tasks. Fig. 1 shows that there were differences in speed performance between the search tools and search tasks. An ANOVA test was therefore performed to establish whether these differences were statistically significant.

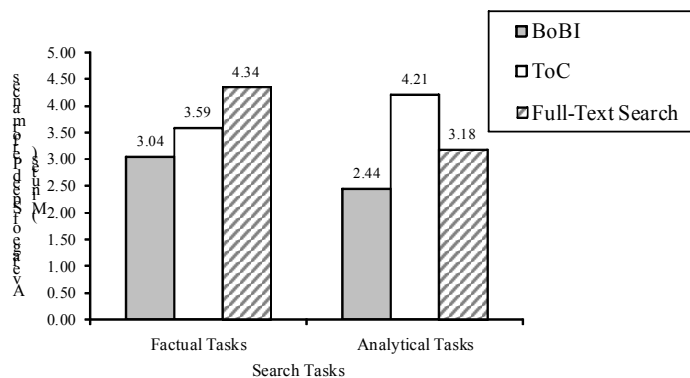


Fig. 1. Speed Performances in Answering Tasks Correctly

Table 1 shows that there was a significant difference in speed performance between the search tools: $F(2,86)=3.22; p<0.04^1$. The search tasks on the other hand exhibited no significant difference: $F(1,43)=0.55; p<0.46$. There was also no significant difference in terms of the Search Tools*Search Tasks interaction: $F(2,86)=0.87; p<0.39$. Hence, it can be concluded that the different types of search task did not significantly affect the speed performance of each of the search tools.

Table 1. ANOVA Test Result for Speed Performance

Independent Variables	Type III Sum of Squares	df	Mean Square	F
Search Tools	276395.73	2	138197.87	3.22
Error (Search Tools)	3689986.93	86	42906.82	
Search Tasks	41350.06	1	41350.06	0.55
Error (Search Tasks)	3215137.61	43	74770.64	
Search Tools * Search Tasks	109259.55	2	75702.73	0.87
Error (Search Tools*Search Tasks)	5386179.78	86	86788.96	

Measure: Speed
Significant at the 0.05 level.

As a significant difference existed in search tool performance, a multiple pairwise comparison test was performed to determine where the difference lay amongst the search tools (at a significance level of 0.05) by comparing it in a pair in order to judge which of the pair has a greater or lower amount of speed performance (minutes). This is shown in Table 2. The table indicated that BoBI and ToC and BoBI and Full-Text Search exhibited significant differences in the search performance ($p<0.04$) in that the BoBI had performed better than the ToC and Full-Text Search. On the other hand, ToC and Full-Text Search showed no significant difference.

Table 2. Multiple Pairwise Comparisons for Search Tools Speed Performance

(I) Search Tools	(J) Search Tools	Mean Difference (I-J)	Std. Error	95% Confidence Interval for Difference	
				Lower Bound	Upper Bound
BoBI	ToC	-73.94	34.93	-144.40	-3.49
	Full-Text Search	-61.68	29.66	-121.50	-1.87
ToC	BoBI	73.94	34.93	3.49	144.40
	Full-Text Search	12.26	28.73	-45.67	70.20
Full-Text Search	BoBI	61.68	29.66	1.87	121.50
	ToC	-12.26	28.73	-70.20	45.67

Measure: Speed (Minutes)
The mean difference is significant at the 0.05 level.

¹ This formula is a standard way of writing an ANOVA test result based on the appropriate table that is referred to - in which "F(df value of the independent variable, df value of error of the independent variable)=F value; $p<Sig.$ (p-value)".

The relative speed differences between the search tools can be seen in Fig. 2 below, where on average the BoBI had the fastest performance at 5.48 minutes, followed by Full-Text Search at 7.52 minutes and finally the ToC at 8.19 minutes. In conclusion, hypotheses H1.1 and H1.2 were validated in that the BoBI was shown to be more efficient when compared to the ToC and Full-Text search tools for performing the search tasks. But there was no conclusive proof that the different search tasks had affected the performance of the search tools.

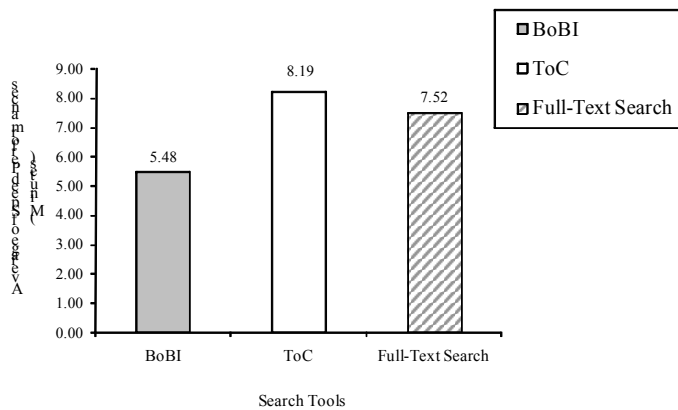


Fig. 2. Overall Speed Performances in Answering Tasks Correctly

4.2 Is a BoBI More Effective Compared to a ToC and Full-Text Search Tool for Finding Information in an e-book Environment?

Fig. 3 below shows the success in finding information accurately in e-books for each search tool based on two types of search tasks. The chart demonstrates that Full-Text Search outperformed the BoBI and ToC for finding information accurately for factual tasks with a success rate of 78%. This was followed by the BoBI (72%) and then the ToC (57%). The BoBI on the other hand, performed the best for accurately finding information in analytical tasks with a 76% success rate. Full-Text Search was second best (69% success) and finally the ToC (67% success). It can be seen from the graph that there were differences in success for finding information accurately amongst the search tools and search tasks in that Full-Text Search was more effective than the BoBI for factual tasks, whereas the BoBI was more effective than Full-Text Search for analytical tasks. These are interesting as a BoBI consists of selective and evaluated entries whereas as Full-Text Search undertakes string matching against a comprehensive index of terms (excluding stop words). An ANOVA test was therefore performed to establish whether these differences were statistically significant.

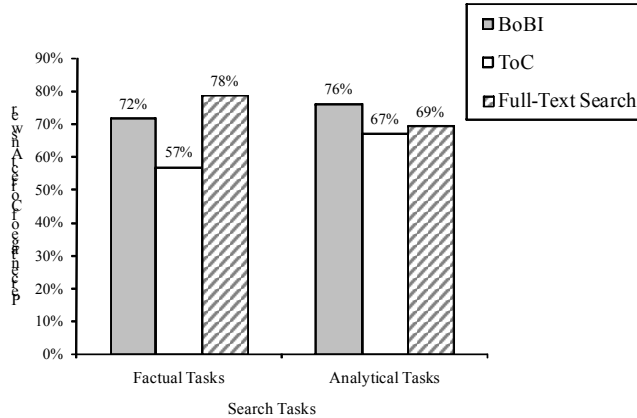


Fig. 3. Count of Success in Finding Information

Table 3 below shows that there were no significant differences using the search tools variable: $F(2,86)=2.60;p<0.08$. This was consistent with the search tasks variable: $F(1,43)=0.30;p<0.59$ and Search Tools*Search Tasks interaction: $F(2,86)=2.14;p<0.12$. Therefore types of search tools, search tasks and their interaction did not have significant differences in terms of the success in finding information accurately. Therefore hypotheses H2.1 and H2.2 were not validated.

Table 3. ANOVA Test Result for Success in Findings Information Accurately

Independent Variables	Type III Sum of Squares	df	Mean Square	F
Search Tools	3.34	2	1.67	2.60
Error(Search Tools)	55.33	86	0.64	
Search Tasks	0.09	1	0.09	0.30
Error(Search Tasks)	13.74	43	0.32	
Search Tools*Search Tasks	1.73	2	0.87	2.14
Error(Search Tools*Search Tasks)	34.93	86	0.41	

Measure: Count of Success

Significant test at the 0.05 level

4.3 Is a BoBI More Useful Compared to a ToC and Full-Text Search Tool for Finding Information in an e-book Environment?

Table 4 below summarises the differences in users' ratings of how useful the three Search Tools were based on the average rank. A Friedman Test was then performed to validate if these differences were statistically significant at the level of 0.05.

Table 4. Users' Rating of the Usefulness of the Search Tools

Search Tools	Average Rank
BoBI	2.27
ToC	1.85
Full-Text Search	1.88

1-4 scale where 1 = useless to 4 = essential

The results of the Friedman test in Table 5 show that there were significant differences in students' ranking of Search Tools usefulness: $X^2(2) = 6.37$, $p < 0.04^2$. A multiple comparison test was then conducted to find the relative differences amongst the Search Tools as shown in Table 6. The test indicated that only the ToC and BoBI had significant differences at the level of 0.05 ($p < 0.02$). It can be concluded that the BOBI had a higher ranking of usefulness (average rank = 2.27) compared to the ToC (average rank = 1.85).

Table 5. Friedman Test of Search Tools Usefulness Rating

Friedman Test	
N	44
Chi-Square	6.37
df	2

* Significant at the 0.05 level

Table 6. Multiple Comparisons for Search Tools Usefulness Rating

Sign Test	Toc - BoBI	Full Text Search - BoBI	Full Text Search - ToC
Z	-2.37346*	-1.54349	0

Measure: Count of Success

* Significant at the 0.05 level

According to the statistical analysis above, it can be concluded that H3.1 was validated in that a BoBI was significantly more useful compared to a ToC for finding information in an e-book environment at 0.05 a significant level. While H3.2 was not validated in that there was not enough evidence to support that the BoBI was more useful compared to a Full-Text Search for finding information in an e-book environment.

4.4 Students' Attitudes towards Search Tools for Finding Information in e-books

4.4.1 Students' Preferences with Respect to a BoBI, ToC and Full-Text Search

Fig. 4 below demonstrates students' ratings of Search Tools in percentages. It can be seen from the graph that most of the respondents rated the BoBI as 'most preferred' (53%) while the ToC was 'second most preferred' (46%) and Full-Text Search was 'least preferred' (34%).

² This formula is a standard of writing Friedman test result based on the appropriate table that is referred to - in which " X^2 (df value of the independent variable)=Chi-Square value; $p < \text{Sig. (p-value)}$ ".

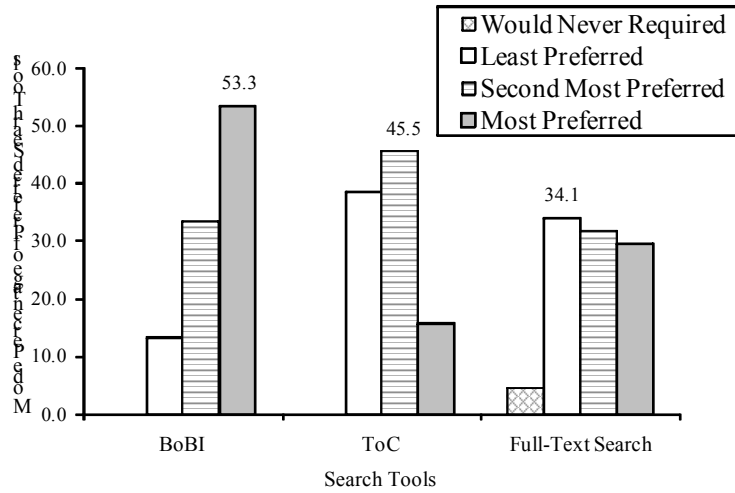


Fig. 4. Users' Rating of Preferred Search Tools

4.4.2 Students' Satisfaction with Respect to a BoBI, ToC and Full-Text Search

Fig. 5 below shows students' rating of how satisfactory Search Tools were in use. It can be seen from the graph that most of the respondents were 'very satisfied' with the BoBI (42%), while most of them were 'satisfied' with the ToC and Full-Text Search at 44% and 38% respectively.

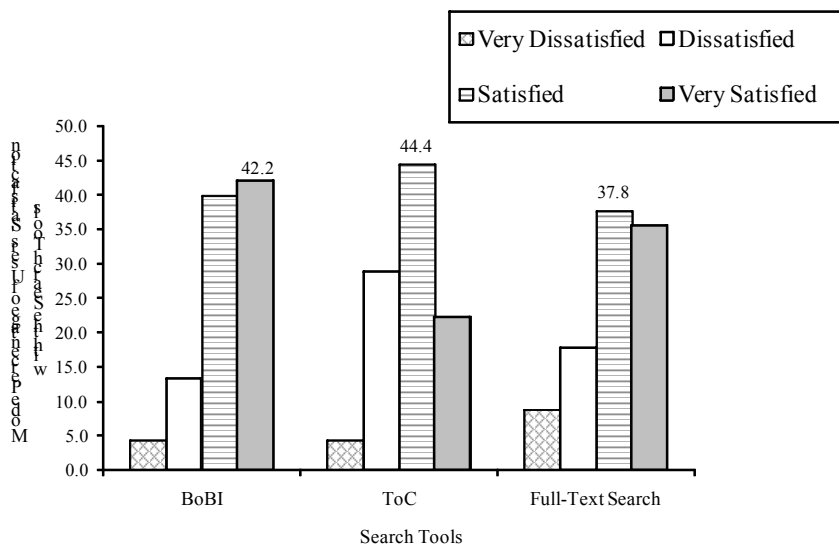


Fig. 5. Users' Rating on Satisfactory with the Search Tools

4.4.3 Students' Rating with Respect to Ease of Use with BoBI, ToC and Full Text Search

Fig. 6 below indicates students' rating of how easy or difficult they found it to use the Search Tools. As can be seen from the graph, most of the respondents found that all the search tools were 'easy' to use but at slightly different levels. The ToC had the highest mode percentage which was 49%, followed by the BoBI (42%), and finally Full-Text Search in 40%.

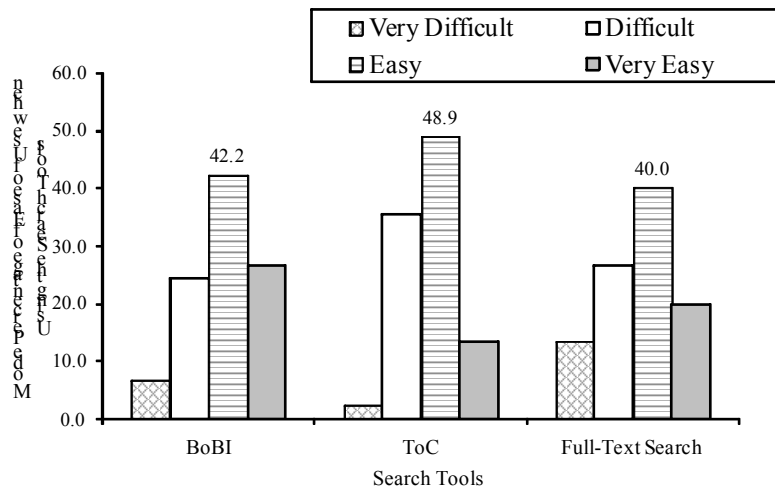


Fig. 6. Users' Rating of Ease of Use when Using the Search Tools

5 Conclusion

This study has demonstrated that a BoBI was significantly more efficient (faster) and useful compared to a ToC and Full-Text Search tool for finding information in an e-book environment. Preference ratings indicated that most students rated the BoBI as 'most preferred' with the ToC as 'second most preferred' and Full-Text Search as 'least preferred'. The search tools were rated similarly for satisfaction and ease of use, but the ToC had the highest mode percentage for satisfaction and ease of use with the BoBI second and finally the Full-Text Search. There was not enough evidence however to support that the different Search Tasks had effectively and efficiently affected the performance of the Search Tools. Although it is not directly comparable in terms of the evaluation design, in the main these findings are inline with [13] and [14].

As a conclusion, the e-book usability evaluation findings are important in gaining a better understanding of the retrieval performance of three search tools (BoBI, ToC and full text search) for browsing for relevant, and searching for specific information in e-books. This will be of value for designing better e-books and access systems.

It is important to acknowledge that the experiment presented here was constrained by subject-specific users and test materials (in information retrieval field) and

therefore the generalisation of the results across other subject fields should be treated with caution.

6 References

- 1 Coyle, K. et al. Report on California Digital Library Joint Steering Committee for shared collections ebook task force. (2001). Available from: <http://www.cdlib.org/inside/groups/jsc/ebooks/>. Last accessed 11/4/2005
- 2 Landoni, M. and Gibb, F. The role of visual rhetoric in the design and production of electronic books: the visual book. *The Electronic Library*, 18(3), pp190-201 (2000)
- 3 Landoni, M., Wilson, R. and Gibb, F. From the Visual Book to the WEB Book: the importance of design. *The Electronic Library*, 18(6), pp407-419 (2000)
- 4 Landoni, M., Crestani, F. and Melucci, M. The Visual Book and the Hyper-TextBook: two electronic books, one lesson? In: *Proceedings of the RIAO 2000 Conference*, Paris, April 12-14, 2000, pp247-265, Paris: CID-CASIS (2000)
- 5 Landoni, M., Wilson, R. and Gibb, F. From the Visual Book to the WEB Book: the importance of design. *The Electronic Library*, 18(6), pp407-419 (2000)
- 6 Wilson, R., Landoni, M. and Gibb, F. A user-centred approach to e-book design. *The Electronic Library*, 20(4), pp322-330 (2002)
- 7 Brown, F. Electronic media and the future of indexing. (1995). Available from: <http://www.allegrotechindexing.com/article01.htm>. Last accessed 20/3/2005
- 8 Bennion, B.C. Performance testing of a book and its index as an information retrieval system. *Journal of the American Society for Information Science*, 31(4), pp264-270 (1980)
- 9 Egan, D.E. et al. Formative Design-Evaluation of SuperBook. *ACM Transactions on Information Systems*, 7(1), pp30-57 (1989)
- 10 Egan D.E. et al. Hypertext for the electronic library? CORE sample results. In: *Proceedings of the 3rd Annual ACM Conference on Hypertext*, San Antonio, Texas, December 15-18, 1991, pp299-312, New York: ACM Press (1991)
- 11 Mynatt B.T. et al. Hypertext or book: which is better for answering questions? In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, Monterey, California, May 3-7, 1992, pp19-25, New York: ACM Press (1992)
- 12 Catenazzi, N. and Sommaruga, L. Hyperbook: an experience in designing and evaluating electronic books. *Journal of Documentation and Text Management*, 2(2), pp81-102 (1994)
- 13 Barnum, C. et al. Index versus full-text search: a usability study of user preference and performance. *Technical Communication*, 51(2), pp185-206 (2004)
- 14 Henke, H. *An Empirical Design for eBooks*. Niwot: Chartula Press (2003)
- 15 Wilson, R. and Landoni, M. Evaluating electronic textbooks: a methodology. In: *Proceedings of the 5th European Conference on Research and Advanced Technology for Digital Libraries*. Darmstadt, September 4-9, 2001, pp1-12, Berlin: Springer (2001)
- 16 Crestani, F. and Ntioudis, S.P. User centred valuation of an automatically constructed Hyper-Textbook. *Journal of Educational Multimedia and Hypermedia*, 11(1), pp3-19 (2001)
- 17 Vakkari, P. Task-based information searching. *Annual Review of Information Science and Technology*, 37, pp413-464 (2003).
- 18 Kinnear, P.R. and Gray, C. D. *SPSS 14 Made Simple*. East Sussex: Psychology Press (2006).