Expandable indexes vs. sequential menus for searching hierarchies on the World Wide Web

PANAYIOTIS ZAPHIRIS†‡, BEN SHNEIDERMAN† and KENT L. NORMAN†

†University of Maryland, College Park, MD 20742, USA

‡The Centre for HCI Design, School of Informatics, City University, London, EC1V 0HB, UK;
e-mail: zaphiri@soi.city.ac.uk

Abstract. An experiment is reported that compared expandable indexes providing full menu context with sequential menus providing only partial context. Menu depth was varied using hierarchies of two, three and four levels deep in an asymmetric structure of 457 root level items. Menus were presented on the World Wide Web within a browser. Participants searched for specific targets. Results suggest that reducing the depth of hierarchies improves performance in terms of speed and search efficiency. Surprisingly, expandable indexes resulted in poorer performance with deeper hierarchies than did sequential menus.

1. Introduction

Information retrieval from the World Wide Web (WWW) is becoming a daily activity in both work and leisure environments. In order to make information retrieval more efficient, it is necessary that indexes, menus and links be carefully designed.

The goal of this study was to investigate the use of expanding hierarchical indexes. One advantage of using such indexes for menu selection is that they preserve the full context of the choice within the hierarchy. While the user browses through the hierarchical structure, the tree is fully displayed. Thus at any point, the user has access to the whole set of major and same level categories. Figure 1 shows a series of displays depicting the expandable index with a menu depth of 4.

Sequential menus, on the other hand, do not display the full hierarchical context as they drop down to deeper levels in the hierarchy. Only elements in the selected category are displayed as options for browsing. This is of particular importance on the WWW when the number of root levels alternatives is large and the depth

of the hierarchy is greater than two. Figure 2 shows an example of the same path through the sequential menu with a depth of 4.

Depth versus breadth in hierarchical menu structures has been the topic of much research. The trade-off between menu depth and breadth is considered by some researchers as the most important aspects that must be considered in the design of hierarchical menu systems (Jacko *et al.* 1994). Miller (1981) found that short-term memory is a limitation of the increased depth of the hierarchy. His experiment examined four structures (64^1, 2^6, 4^3 and 8^2) with a fixed number of target items (64). As depth increased so did the response time to select the desired item.

Snowberry *et al.* (1983) replicated Miller's study by examining the same structures but this time including an initial screening session during which participants took memory span and visual scanning tests. They found that instead of memory span, visual scanning was predictive of performance, especially in the deepest hierarchies.

Kiger (1984) extended Miller's research by doing an experiment that provided users with five modes of varying menu designs of 64 end nodes (2^6, 4^3, 8^2 and 16x4, 4x16). Performance and preference data were collected. The results of the experiment showed that the time and number of errors increased with the depth of the menu structure. The 4 x 16 structure had the fastest response times and the fewest errors. The participants ranked the menus with least depth as the most favourable (the 8^2 structure was favoured).

An experiment by Jacko and Salvendy (1996) tested six structures (2^2, 2^3, 2^6, 8^2, 8^3 and 8^6) for reaction time, error rates and subjective preference. They demonstrated that as depth of a computerized, hierarchical
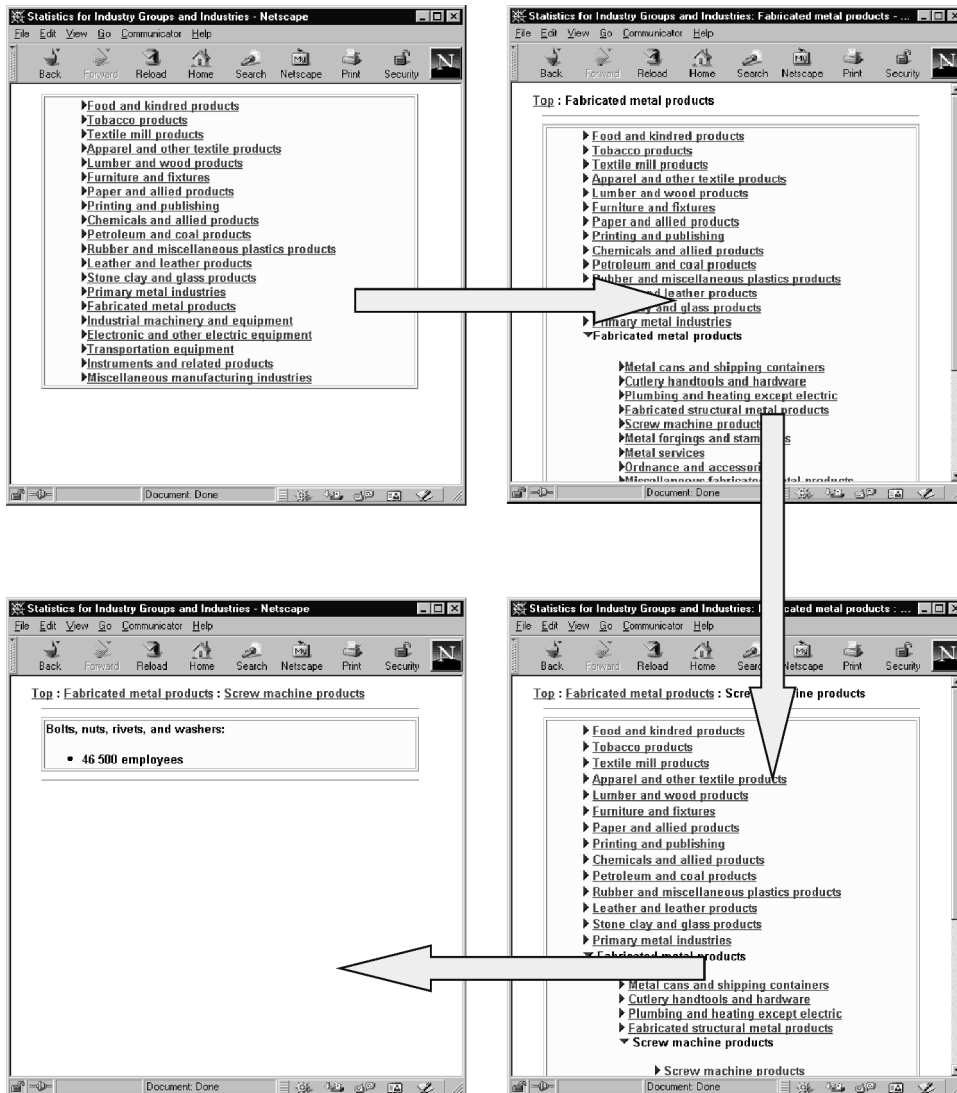
Figure 1.   Expandable index menu with hierarchical depth of 4. Arrows show the series of selections.

menu increased, perceived complexity of the menu increased significantly. Campbell (1988) identified multiple paths, multiple outcomes, conflicting interdependence among paths and uncertain linkages as four characteristics of a complex task. Jacko and Salvendy build on this framework to suggest that these four characteristics are present as depth increases, and the presence of these four characteristics is responsible for the increase in complexity.

Wallace *et al.* (1987) confirmed that broader, shallower trees (4 x 3 versus 2 x 6) produced superior performance, and showed that, when users were stressed, they made 96% more errors and took 16% longer. The stressor was simply an instruction to work quickly ('It is imperative that you finish the task just as quickly as possible'). The control group received mild instructions to avoid rushing ('take your time; there is no rush').

Norman and Chin (1988) fixed the number of levels at four, with 256 target items, and varied the shape of the tree structures. They recommend greater breadth at the root and at the leaves, and added a further encouragement to minimize the total number of menu frames needed so as to increase familiarity.

Zaphiris and Mtei (1997) replicated Kiger's (1984) structures but this time on the WWW using hyperlinks. Overall, their results were in agreement with those of Kiger (1984). They found that of the structures tested (2^6, 4^3, 8^2 and 16 x 4, 4 x 16), the 8^2 structure was the fastest to search.

Larson and Czerwinski (1998) carried out an experiment using 512 bottom level nodes arranged in three
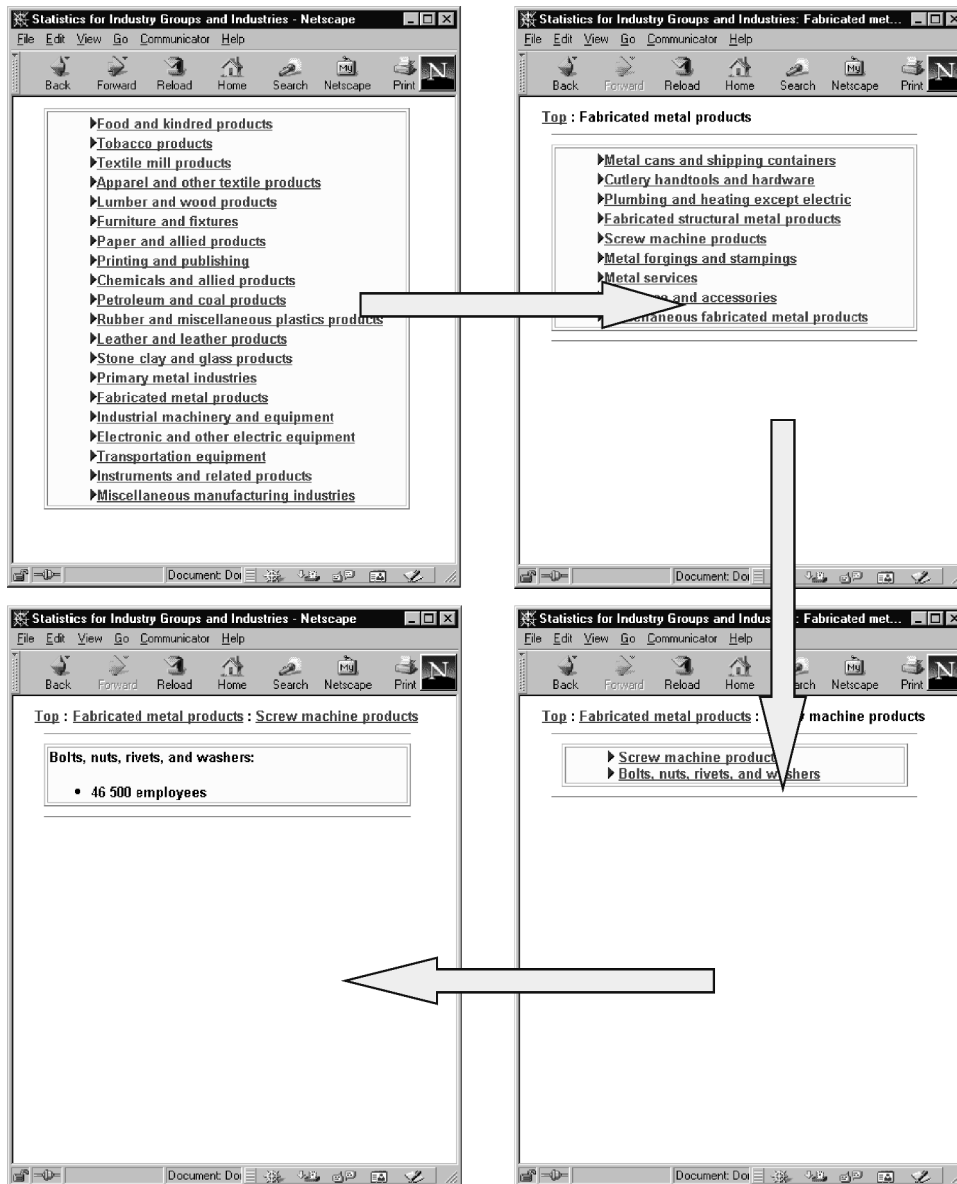
Figure 2. Sequential menu with hierarchical depth of 4. Arrows show the series of selections.

different structures (8 x 8 x 8, 32 x 16, 16 x 32). Participants on average completed search tasks faster in the 16 x 32 hierarchy, second fastest in the 32 x 16 hierarchy, and slowest in the 8 x 8 x 8 hierarchy. In addition, on average, participants tended to be lost least often in the 16 x 32 hierarchy. Larson and Czerwinski calculated 'lostness' through an analysis of the number of unique and total links visited in comparison to the 'optimal' path.

One limitation with previous studies on the topic of menu selection has been the use of symmetric hierarchical structures that do not reflect typical menu structures in real-world applications. Based on a more ecologically valid approach, the present experiment involves an asymmetric hierarchical structure that is common in most practical applications (Norman 1990).

While previous research has investigated the depth versus breadth effect, it has neglected the effect of type of menu presentation. Specifically, the comparison between expandable indexes with full context and sequential menus with limited context in hierarchical structures on the WWW has not received sufficient attention despite the fact that many designers are employing expandable indexes in current WWW design. It was predicted that expandable indexes would result in superior performance because they provide a full

context for the choices at lower levels of the hierarchy. Furthermore, sequential menus are predicted to be less efficient and to result in more backtracking due to loss of context and forgetting on the part of the user.

2. Method

2.1. *Participants*

Twenty-one students (14 males and 7 females) at the University of Maryland with at least an undergraduate degree in business and management took part in this study. The mean age of the participants was 27. They were all experienced computer and WWW users who had been using the WWW for at least one year and at least twice per week. All participated voluntarily signed a consent form after agreeing to the terms of the experiment.

2.2. *Materials*

Menu hierarchies were constructed and hosted on the WWW. The bottom level nodes for the various hierarchical structures were taken from the 1996 Annual Survey of Manufacturers 'Statistics for Industry Groups and Industries' of the United States Bureau of the Census. This survey consisted of a total of 457 bottom level nodes arranged in a specific hierarchy. This structure was used as the basis for designing the different WWW menu organizations used in this study.

Menu structures were generated that varied along two factors: depth of the hierarchy and method of display. Three levels of depth were used: 2, 3 and 4. The method of display used was either expandable indexes (Expandable) or sequential menus (Sequential). A factorial design (3 x 2) was used to generate all six possible combinations of levels.

Pages were generated for display in a WWW browser (Netscape 4.0). As shown in figures 1 and 2, the first level of the hierarchy contained the top level choices only, the intermediate hierarchies contained the different options and a top hyperlink navigation bar that allowed users to move back to previous pages or to return to the top level of the menu hierarchy and the final level node contained the search target information (the number of employees in the specific industry) together with the top hyperlink navigation bar.

2.3. *Procedures*

Each participant was first told the purpose of the experiment and was then presented with the basic instructions to follow throughout the experiment. Specifically, they were told that we were interested in determining the optimal depth in presenting links on the WWW for expandable indexes and sequential menus that varied the amount of context. In addition, they were given the following general instructions:

- They were to browse through the assigned pages until they felt comfortable that they found the answer to the assigned task (the task was to locate the number of employees in a specific industry) questions.
- They were to perform all searches as quickly as possible while making as few mouse clicks as possible.
- They were to navigate through the web pages using only the links present in the pages and not the navigation buttons of the browser.
- They were instructed not to browse for any other information that they have not been asked to seek until they had finished all assigned tasks.
- They should read carefully the assigned tasks and instructions presented to them.

After explaining to the participants their tasks, they were given the opportunity to ask questions. All those that decided to participate were asked to read and sign a consent form.

Participants were divided into three randomly selected groups (between subjects factor) and assigned to one of the three levels of menu depth. Menu type (Expandable or Sequential) was a within subjects factor and was counterbalanced by presenting different treatment orders. Half of the participants in each group worked with the Expandable Menu first and the Sequential Menu second. The remaining half worked with Sequential Menu first and Expandable Menu second. Each participant performed a total of 10 searches (five in a Expandable Menu and five in a Sequential Menu). The browser height was set to just display 20 lines of text. Each trial's target and hierarchy information was presented to the participant on paper with one target presented per page next to the computer so it was constantly available as a reference. At the start of each trial, the participant was asked to turn the page in order to see the new target and menu type to be used for the next trial.

Three kinds of data were collected for analysis: response times, search efficiency and subjective ratings. Response times and retrieval efficiency data were collected from the WWW server log. Response time was calculated as the time taken by the participant to browse from the main page to the specific bottom level node. Efficiency of retrieval was calculated as the

difference in mouse clicks between the total number of mouse clicks to reach the bottom level node minus the 'optimal' number of mouse clicks required to reach the target bottom level node. After the participants finished all 10 tasks assigned, subjective preference responses were collected for the Expandable and Sequential Menus by answering on a scale of 1 to 9 the following three subjective questions:

- which of the two types of menus browsed would you prefer to use?;
- rate the ease of navigation;
- rate your sense of orientation.

Each experimental session lasted around 20 minutes.

3. Results

3.1. *Response times of information retrieval*

Figure 3 shows that with the Expandable Menu participants completed search tasks fastest in the hierarchy of Depth 2, second fastest in the Depth 3, and slowest in Depth 4. Menu depth had less of an effect for the Sequential Menus. The average response time for Depths 2 and 3 was about the same but faster than the response time for Depth 4. Surprisingly, searches were slower for the Expandable Menus than for the Sequential Menus, particularly for hierarchy depths of 3 and 4.

The mean response times and corresponding standard deviations for each hierarchy are shown in table 1. A split-plot analysis of variance indicated that the main effect of Menu Type was significant ($F(1, 18) = 13.91$, p

Table 1. Mean response times for Menu Type (Expandable or Sequential) and Menu Depth (2,3 or 4) (standard deviations are in parentheses).

	Expandable			Sequential		
Depth	2	3	4	2	3	4
Response time (seconds)	39.6 (36.4)	46.7 (27.7)	65.8 (53.2)	35.6 (30.3)	34.9 (19.1)	41.1 (39.8)

<0.01). On the average, response times were nearly 50% longer for the Expandable Menus than for the Sequential Menus. In addition, there was a significant interaction between Menu Type and Menu Depth ($F(2, 18) = 3.62$, p <0.05). The effect of Menu Depth was greater for the Expandable Menus than for the Sequential Menus. Finally, the main effect of Menu Depth was not significant ($F(2, 18) = 1.34$, p >0.05). This was probably due to the low statistical power associated with the between-subject factor.

3.2. *Efficiency of information retrieval*

An analysis of the number of mouse clicks made above and beyond the most direct path to the target was performed. Search inefficiency is equal to the total number of mouse clicks to complete each task minus the 'optimal' minimum number of clicks. On the average, participants in the Depth 2 hierarchies made fewer mouse clicks in the Expandable Menu than in the Sequential Menu. For Depth 3 hierarchies, performance was about the same for the two menu types. However, in the Depth 4 hierarchies performance was drastically worse with the Expandable Menu than with the Sequential Menu. The mean number of mouse clicks beyond the most efficient path for each hierarchy are graphed in figure 4 and listed in table 2 along with corresponding standard deviations.

A split-plot analysis of variance revealed a significant interaction between Menu Depth and Menu Type ($F(2, 18) = 4.75$, p <0.05). The effect of Menu Depth was much larger for the Expandable Menus than for the Sequential Menus. On the other hand, the main effects of Menu Depth and Menu Type were not significant ($F(2, 18) = 1.97$, p >0.05 and $F(1, 18) = 1.34$, p >0.05, respectively) owing to the form of the interaction.

3.3. *Subjective ratings and preference measures*

After finishing the search tasks, each participant completed a subjective satisfaction questionnaire composed of three ratings: preference to use, ease of
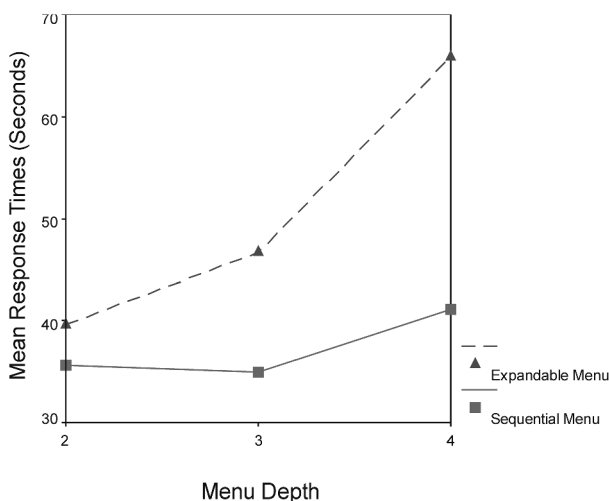

Figure 3. Mean response times as a function of Menu Type (Expandable or Sequential) and Menu Depth (2,3 or 4).

navigation and sense of orientation. Table 3 lists the preference ratings (means and standard deviations) for the Expandable Menu over the Sequential Menu. These results indicate that participants had a slight preference for the Expandable Menus for depths of 2 and 3 but for depth 4 their preference is to use the Sequential Menu. However, none of these effects were significant $(F(5,34) = 0.57, p > 0.05)$.
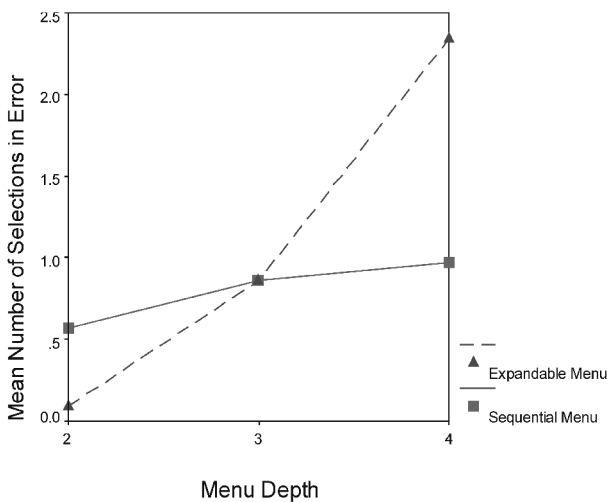

Figure 4. Inefficiency of information retrieval as a function of Menu Type (Expandable or Sequential) and Menu Depth (2,3 or 4).

Table 2. Inefficiency of information retrieval (standard deviations are shown in parentheses).

	Expandable			Sequential		
Depth	2	3	4	2	3	4
Mouse clicks in error	0.09 (0.28)	0.86 (2.12)	2.34 (4.30)	0.57 (1.14)	0.86 (1.52)	0.97 (2.56)

Table 3. Mean preference ratings (1 = preference for Sequential Menu. 9 = preference for Expandable Menu), mean ratings of ease of navigation (1 = low. 9 = high) and mean ratings of sense of orientation (1 = low. 9 = high).

	Expandable			Sequential		
Depth	2	3	4	2	3	4
Preference ratings	6.00 (1.83)	6.43 (1.51)	4.86 (2.48)	5.14 (2.12)	5.57 (2.44)	6.29 (2.14)
Ease of navigation	5.71 (1.25)	6.29 (1.25)	5.57 (1.91)	6.29 (1.50)	6.14 (1.35)	6.29 (1.38)
Sense of orientation	6.00 (1.73)	6.29 (1.70)	5.00 (1.73)	6.71 (1.38)	6.29 (1.38)	6.29 (1.80)

Results are from the replies of 21 participants with standard deviations in parentheses.

Table 3 also lists the ease of navigation ratings (means and standard deviations) for the Expandable and Sequential menus. Participants tended to rate the Sequential Menus as slightly easier to navigate than the Expandable Menus for Depths of 2 and 4 but about equal for Depth of 3. However, again, the analysis of variance test indicated no significant effect due to ease of navigation $(F(5,34) = 0.33, p > 0.05)$.

Finally, when asked to rate their sense of orientation, participants rated the Sequential Menu as providing them with a slightly better sense of orientation when working with the Depth 2 and 4 hierarchies than the Expandable Menu. However the analysis of variance test indicated no significant main effect with respect to sense of orientation $(F(5,34) = 1.04, p > 0.05)$.

4. Discussion

The response time and efficiency of retrieval data support previous research on the issue of depth vs. breadth in hierarchical menu structures. Performance—both in terms of access time and efficiency—decreases as the depth of the menu structure increases. The present results extend this conclusion to asymmetric menus as well.

It was expected that menus employing the expandable indexes providing the full context of choices within the hierarchy would result in better performance than sequential menus that provide only limited context. However, this was not the case, particularly for deeper hierarchies where one would expect the context information to be even more beneficial. Surprisingly, tasks were completed significantly faster with Sequential Menus than with Expandable Menus. Efficiency of search data indicated a slightly different pattern, which may help to explain the unexpected result for response time. For Depth 2 hierarchies, the Sequential Menus resulted in slightly more search moves than the Expandable Menu; for Depth 3 performance was equal; but for Depth 4 the Expandable Menu resulted in significantly more search moves than the Sequential Menus.

A possible reason for the poor performance of the Expandable Menu has to do with the problem of long vertical lists on a screen showing only 20 lines of text. When the Expandable Menus were expanded in the Depth 4 hierarchies, they were very long and unwieldy. Users opened and closed more levels of the hierarchy and took longer to find the targets because the long indented indexes were hard to scan and scroll for context information. Furthermore, context may have been lost as the expanded index scrolled out of view in the limited size of the browser window (although some

users tended to use quicker scrolling by clicking the middle of the scroll bar, it was observed that participants needed between zero and six extra clicks on the scroll bar arrow to reach the target link).

User preferences tended to agree with user performance; however, none of these effects were statistically significant. Users tended to prefer to use Expandable Menus for depths of 2 and 3 but when the depth is increased to 4, they chose Sequential Menus as more desirable. In addition, users found the Expandable Menus difficult to navigate (except for Depth 3 hierarchies) and lost their orientation while browsing Expandable Menus (except for Depth 3 hierarchy where Expandable and Sequential Menus had equal ratings for sense of orientation).

4.1. *Implications for WWW designers*

The results of this experiment replicate results of previous research in the area of depth vs. breadth tradeoffs in menu selection. Menu hierarchies should be designed with a minimum depth and maximum breadth if at all possible. Expandable index menus are acceptable only for shallow menu hierarchies or Depth 2 and 3, and should be avoided for deeper hierarchies. If menus with expandable indexes are used, they should be redesigned in such a way as to make relevant, hierarchical, context information clear and available to the user. These designs should avoid hard to follow indentation schemes and long lists that require excessive scrolling of the list in a browser window.

4.2. *Suggestions for future research*

Additional research is required to find out precisely why the expandable indexes resulted in such poor performance with deeper hierarchical menus. In particular, what is the problem with showing context using hierarchical indentation? How can the problem of scrolling long lists be avoided perhaps by expanding the hierarchies in a horizontal direction? Is there some ideal combination of sequential menus and expandable menus that could improve user performance?

References

CAMPBELL, D. J. 1988, Task complexity: a review and analysis. *Academy of Management Review*, **13**, 40–52.

JACKO, J. and SALVENDY, G. 1996, Hierarchical menu design: breadth, depth, and task complexity. *Perceptual and Motor Skills*, **82**, 1187–1201.

JACKO, J., KOUBEK, R., SALVENDY, G. 1994, An experimental study for menu design: Guidelines for menu selection. *Proceedings of the 12th Triennial Congress of the International Ergonomics Association*, Toronto, Canada, August 15–19, **4,** 351–353.

KIGER, J. I. 1984, The depth/breadth tradeoff in the design of menu-driven interfaces. *International Journal of Man-Machine Studies*, **20**, 201–213.

LARSON, K. and CZERWINSKI, M. 1998, Page design: implications of memory, structure and scent for information retrieval. In the *Proceedings of CHI '98 Human Factors in Computing Systems*, Los Angeles, CA, April, 1998, ACM Press, pp. 25–32.

MILLER, D. P. 1981, The depth/breadth tradeoff in hierarchical computer menus. Proceedings of the *Human Factors Society*, 296–300.

NORMAN, K. 1990, *The Psychology of Menu Selection: Designing Cognitive Control of the Human/Computer Interface* (Norwood, NJ: Ablex Publishing Corporation).

NORMAN, K. and CHIN, J. 1988, The effect of tree structures on search in a hierarchical menu selection system. *Behaviour and Information Technology*, **7**, 51–65.

SNOWBERRY, K., PARKINSON, S. R. and SISSON, N. 1983, Computer display menus. *Ergonomics*, **26**, 699–712.

WALLACE, D., ANDERSON, N. and SHNEIDERMAN, B. 1987, Time stress effects on two menu selection systems. *Proceedings of the Human Factors Society, Thirty-First Annual Meeting (1987)*, pp. 727–731.

ZAPHIRIS, P. and MTEI, L. 1997, Depth vs. Breadth in the Arrangement of Web Links. Available online: http://otal.umd.edu/SHORE/bs04/