Age Related Differences and the Depth vs. Breadth Tradeoff in Hierarchical Online Information Systems

Panayiotis Zaphiris[1], Sri Hastuti Kurniawan[2], and R. Darin Ellis[3]

[1]The Centre for HCI Design, City University
London, EC1V 0HB, U.K.
zaphiri@soi.city.ac.uk
[2]Department of Computation, UMIST
Manchester, U.K.
s.kurniawan@co.umist.ac.uk
[3]Institute of Gerontology, Wayne State University
Detroit, MI 48202, USA

Abstract. We report the results of a study investigating the age related differences as they relate to the depth versus breadth tradeoff in hierarchical online information systems. Different stimulus of various depth and breadth combinations and expandable or non-expandable structure were used. Participants from two age groups (aged 36 years old and younger or 57 years old and older) took part in this study. Overall, shallow hierarchies were preferred to deep hierarchies. Seniors were slower but did not make more errors than their younger counterparts when browsing the different treatments.

1 Introduction

1.1 Defining Information Architecture

Information architecture has been defined as the "process of structuring and organizing information so that it is easier for users to find and for owners to maintain" [1] and as "a structure or map of information which allows others to find their personal paths to knowledge" [2] or as "simply a set of aids that match user needs with information resources" [3].

Although reports [4] stated that roughly two thirds of users browsing the World Wide Web (WWW) are looking for specific information and that people do not come to the web for an "experience" but rather for the information [5], the impact of a web site's information architecture on the ability of a user to navigate is very often overlooked by many web site designers.

How information is categorized, labeled and presented and how navigation and access are facilitated determines not only whether users will and can find what they need, but also affect user satisfaction and do influence return visits [6].

How well the information architecture of a site is designed also effects users' sense of orientation (knowing where they are in the hierarchy). The problem of disorientation in hypertext generally results in a measurable decline in user performance [7]. The disoriented user is usually unable to gain an overview of the

N. Carbonell, C. Stephanidis (Eds.): User Interfaces for All, LNCS 2615, pp. 23–42, 2003.

material, encounters problems in deciding if needed information is available, and has problems in deciding where to look for information and how to get there [8, 9]. Furthermore, disorientation can lead users to repeatedly open the same few nodes [10] and can increase the time users take to locate information, due to following less than optimal routes through the hypertext [11].

In general, most people need to know some basic information in order to orient themselves. This is true whether they are navigating a building or a web site. These basic navigation needs include answers to the following questions [12]:

- Where am I?
- Where can I go?
- How can I get back to where I once was?

Hierarchical structuring is one of the best ways to present information, especially to help non-knowledgeable users in navigating [11]. Two variations of the standard hierarchical structure is the expandable and non-expandable hierarchies or expandable indexes and sequential menus as they are respectively mentioned in Zaphiris, Shneiderman and Norman [13].

One advantage of using expandable hierarchies 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.

Non-expandable hierarchies, 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. Examples of expandable and non-expandable hierarchies are depicted in Fig. 1 and Fig. 2.

1.2 Depth versus Breadth in Menu Selection

Menu panels usually consist of a list of options. These options may consist of words or icons. The word or icon conveys some information about the consequences of selecting that option. Sometimes the options are elaborated with verbal descriptors.

When one of the options is selected and executed, a system action occurs that usually results in a visual change on the system display. The total set of options is usually distributed over many different menu panels/frames. Web indexes are organized in a similar structure. Links (very often 2-3 words, sometimes elaborated with verbal descriptors) are arranged in various levels of homepages. These links convey information about the page (with information or further sub-categories) that will be displayed if that specific link is selected.

The topic of menu selection and especially the depth versus breadth tradeoff (Examples of deep and shallow web hierarchies are shown in Fig. 3 and Fig. 4.) has been extensively examined, both empirically and analytically. The navigation problem (i.e., getting lost or using an inefficient pathway to the goal) becomes more and more treacherous as the depth of the hierarchy increases. Research has shown [14] that error rates increased from 4.0% to 34.0% as depth increased from a single level to six levels.

It has been shown experimentally [15, 16] that hierarchical menu design experiments can be replicated to experiments on hierarchies of lists of web links. Hierarchical decomposition of the user's selection of action is often necessary to facilitate fast and accurate completion of search tasks, especially when there is insufficient screen space to display all possible courses of action to the user.

The challenge, therefore, is to enable the user to select the desired course of action using a clear, well-defined sequence of steps to complete a given task [17].

1.3 Empirical Results

Depth versus breadth in hierarchical menu structures has been the topic of much empirical research. The trade-off between menu depth (levels in a hierarchical menu structure) and breadth (number of menu selections present on a given menu in the hierarchy) is stated by some researchers as the most important aspect that must be considered in the design of hierarchical menu systems [18]. Questions such as "how long does it take to find a specific piece of information?" and "how many errors occur during the information search?" have been the main focus of previous empirical experimental research on the topic of menu selection.

Miller [19] found that short-term memory is a limitation of the increased depth of the hierarchy. His experiment examined four structures (64^1, 8^2, 4^3, and 2^6) with a fixed number of target items (64). The 64 items were carefully chosen so that "they form valid semantic hierarchies" in each of the 4 menu structures. As depth increased, so did response time to select the desired item.

Snowberry, Parkinson and Sisson [14] replicated Miller's study by examining the same structures, but this time included an initial screening session during which subjects took memory span and visual scanning tests. They stated that determination of whether short-term memory and visual scanning act as limiting factors in menu selection is also important for consideration of 'special' groups, such as older users, whose short-term memory capacity and processing speed are reduced relative to young adults. They found that instead of memory span, visual scanning was predictive of performance, especially in the deepest hierarchies. Their experiment compared performances on both randomized and categorized 64 item displays to performances on structures of increasing depth and decreasing breadth. They used a between-subject design with four different experimental conditions (menu structures - 64^1, 2^6, 4^3 and 8^2).

Their results provide a nice replication of Miller's results showing that search time for randomly arranged items for the four experimental groups produced a U-shaped function, with a minimum at the configuration of two levels with eight choices. With category organization held constant across conditions, search time improved as depth decreased and breadth increased. Accuracy was found to improve as breadth increased (and depth decreased) regardless of display format in the broadest display. They state three possible kinds of 'forgetting' which might have led to lower accuracy for deeper menus:

1. Participants might have forgotten the target word.
2. Participants might have performed less accurately in deep menus because they forgot the pathway to the target.

3. Rather than associating a target with a path of options, participants might have based selections of options purely on the association formed between the items displayed and the target; forgetting the association would have led to inaccurate performance.

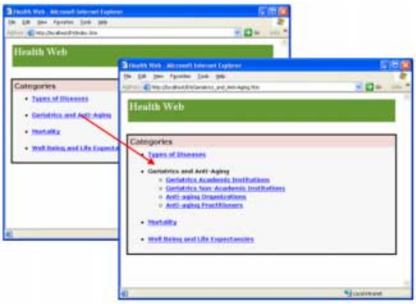

Fig. 1. Expandable hierarchical structure

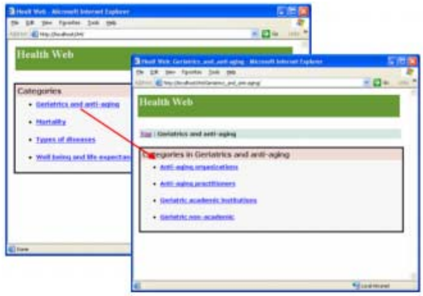

Fig. 2. Non-expandable hierarchical structure

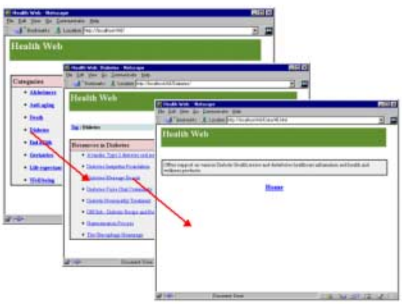

Fig. 3. Broad hierarchical web index

Fig. 4. Deep hierarchical web index

Kiger [20] extended Miller's research by designing an experiment that provided users with five modes of varying menu designs of 64 end nodes.

- 8^2: 8 items on each of 2 levels.
- 4^3: 4 items on each of 3 levels.
- 2^6: 2 items on each of 6 levels.
- $4^1 + 16^1$: A 4 item menu followed by a 16 item menu
- 16^1+4^1: A 16 item menu followed by a 4 item menu

Kiger's experiment [20] showed that the time and number of errors increased with the depth of the menu structure. The 4x16 structure had the fastest response times and the fewest errors. The participants ranked the menus with least depth (the 8^2 structure) as the most favorable.

Parkinson et al. [21] built on Snowberry, Parkinson and Sisson's [14] research by examining different organization styles of menus. Three factors were varied in the categorized menus they used: alphabetical versus categorical ordering of words within categories; spacing versus no additional spacing between category groups; and category organization arranged by column or by row. Organization (row or column) and spacing exerted powerful effects on search time. Menus in which categories were arranged by column were searched faster than menus organized by row. Menus in which additional spacing was provided between groups were searched faster than menus with no spacing. For categorized menus, no advantage was found for alphabetical as opposed to categorical ordering of words within category. These results are in disagreement with previous research by Mc Donald et al. [22] who found that categorical menu organization is superior to pure alphabetical or random organization.

Similar results were also found in an experiment by Nygren [23] where scanning of a horizontal listing of items was found to be slower than scanning a vertical listing of items. Findings further indicated that scanning a single long vertical list was faster than scanning multiple shorter vertical lists. They also found that estimated scanning rate was dependent on subject's age. Searching for a target item was significantly faster if the target item was given a unique feature, compared with if no unique feature was used. The color, shade, space, slant and size features were equally effective compared with a control condition (without features).

Wallace et al. [17] confirmed that broader, shallower trees (4x3 versus 2x6) produced superior performance, and showed that, when users were stressed, they made 96 percent more errors and took 16 percent more time to browse to the target. 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 [24] 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.

An experiment by Jacko and Salvendy [18] tested six structures (2^2, 2^3, 2^6, 8^2, 8^3, and 8^6) for search time, error rates, and subjective preference. They demonstrated that as depth of a computerized, hierarchical menu increased, perceived complexity of the menu increased significantly. Campbell [25] identified multiple paths, multiple outcomes, conflicting interdependence among paths, and uncertain linkages as four characteristics of a complex task. Jacko and Salvendy [18] built 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.

Studies have also shown [26] that apart from the inherent complexity of a hierarchical structure, the depth versus breadth tradeoff is also influenced by cases where the user can restrict the scope of search either because of experience or because the options are organized into categories.

As mentioned earlier, the similarity between menu panels and hierarchical lists of web links has recently been studies (through a series of replication of previous web panel experiments) by researchers.

Zaphiris [16] replicated Kiger's [20] structures but this time on the WWW using hyperlinks. Overall, his results were in agreement with those of Kiger [20]. He found that of the structures tested (8^2, 4^3, 2^6 and 16x4, 4x16), the 8^2 structure was the fastest to search.

Larson and Czerwinski [15] carried out an experiment using 512 bottom level nodes arranged in three different structures (8x8x8, 32x16, 16x32). Subjects on average completed search tasks faster in the 16x32 hierarchy, second fastest in the 32x16 hierarchy, and slowest in the 8x8x8 hierarchy. Also, on average, subjects tended to be lost least often in the 16x32 hierarchy. The authors calculated "lostness" through an analysis of the number of unique and total links visited in comparison to the "optimal" path. Similar results to the above two web related studies have been found in a recent study by Zaphiris, et al. [13] when investigating the difference between sequential and expandable hierarchies of links.

In a study by Mead, et al. [27], age-related differences and training on WWW navigation strategies were examined. They found that novice users were more likely to "get lost" in a hierarchical structure than were more experienced searchers. In an experiment they conducted, where they asked their participants to complete nine search tasks on the WWW, they showed that seniors were significantly less likely to complete all tasks than were younger adults. However, seniors were as likely to complete the five tasks with short optimal path length (two moves or fewer) as younger adults. Seniors on the other hand, were significantly less likely to complete the tasks with long optimal path length (3 moves or more) than were younger adults. Error results showed seniors may experience considerable difficulty finding information on the WWW when path length is long. Finally seniors were reported to adopt less efficient search strategies and had more problems remembering which pages they had visited (and what was on those pages) than did younger adults.

1.4 Seniors and the Web

Administration on Aging [28] projected that by the year 2030 people aged 65 years and above will represent 20% of US population.

Recent data shows that among the 21.8 million U.S. households where the householder was 65 or older, 5.3 million households (24.3%) had a computer, slightly less than half the figure for the general population (51.6%). Internet access was present in 3.9 million (17.7%) of the elderly households, whereas for the general population this was 41.5%. On an individual basis, 9.3 million (28.4%) of the older persons had home computer access and 4.2 million (12.8%) used the Internet at home, which is 34% of the figure for the general population [28]. One of the main reasons senior citizens report for using a computer is to look for health information [29, 30].

This increase in computer and web usage by senior citizens is expected to follow a similar patter in other countries too. Various market research showed that the younger generation is getting more and more dependent on the internet/web in performing daily activities (entertainment, shopping, studying). Hence, in 20-30 years, we will be more likely to experience a more dramatic increase in web usage by seniors, making the study of aging issues effecting web usage a necessity, especially since some findings suggested that older adults have some disadvantages in fully utilizing the Internet as an information source. That is, older people have more trouble finding information in a Web site than younger people [27, 31].

One way to alleviate age-related barriers in using the Internet is by involving seniors in the design process. Zaphiris and Kurniawan [32] proposed the term "senior-centered design" to refer to any design methodology that involves older users in the process of designing products that are targeted towards the aging population.

2 Research Methodology

2.1 Stimulus Material

Two sets of locally hosted (to eliminate delays due to network congestions) web sites, each containing 64 pages of health and medical-related information for older adults, were designed and presented to participants using a standard web browser. The test stimuli were taken from Health:Aging and Health:Senior Health directories of dmoz (http://www.dmoz.org). These sites were supposedly categorized by experts in the relevant areas. There were two groups of information used for test stimuli. There were two tested modes of hierarchical information structures: expandable and non-expandable as depicted in Fig. 1 and Fig. 2. There were 12 experimental conditions (two sets of hierarchies, with 2 structures and 3 depth conditions per structure) used as stimuli. For each hierarchical structure and each information group, there were 3 conditions with varying depth/breadth as listed in Table 1.

Table 1. Depth/Breadth combinations used in the stimuli

Depth	2	3	6
Breadth	8	4	2

2.2 Participants

Participants in the general browsing experiment consisted of 24 older Internet users (aged 57 years old and older) and 24 younger users (aged 36 years old and younger), who lived independently in the community (non-institutionalized) with no visual or cognitive impairment or functional illiteracy or any other deficit that would preclude them from successfully completing the protocol.

Participants were recruited from senior centers, senior church groups, community organizations, and the Wayne State University campus. The sample consisted of a variety of ethnic groups, gender, and education levels.

Participants were initially screened through brief conversation with the experimenter to assess their computer and Internet usage. By asking several Web-related questions, individuals with no prior experience with the WWW or who had not used a computer before were screened-out.

Individuals reporting a medical condition that would directly interfere with performance (e.g. severe carpal tunnel syndrome) were also excluded from the study. Participants' demographics data are listed in Table 2.

Table 2. Demographic data of participants of the general browsing experiment

Socio-demographic data		Young	Old
Age		26.8(5.18)*	67.5(6.5)*
Education		16.8(2.57)*	14.6(3.13)*
Gender	Male	18	10
	Female	6	14
Ethnicity	Asian/Pacific Islander	11	0
	African/African American	4	12
	Caucasian/Middle Eastern	9	12

* mean(standard deviation)

2.3 Apparatus

Pentium II-class personal computers operating with Windows 98, SVGA monitors (1280x1024 resolution, .22mm dot pitch @ 85 Hz) were used throughout this experiment. User behavior was recorded using server logs capable of recording time and page clicks. To verify that the server's log recorded the correct time and duration, Lotus Screen Cam® software was used to observe the first few participants.

2.4 Procedure

2.4.1 Pre-experiment Questionnaire

The participants were first asked to read and sign a consent form. The consent form outlined briefly the purpose of the experiment and listed all the participants rights. After signing the consent form, the participants were asked to fill a socio-demographic questionnaire that asked for their age, ethnicity and education level.

Following that, participants were asked to fill two questionnaires. In the first one their general computer and internet knowledge was assessed using a modified questionnaire from the user demographics section of the WWW User Survey of GVU Center of the Georgia Institute of Technology (http://www.gvu.gatech.edu/). In the second questionnaire the participants' domain (health information) pre-experiment knowledge was assessed.

Previous studies suggested that domain knowledge might contribute to the success of information search. Therefore, a bivariate correlation between domain knowledge and search effectiveness was calculated. The result showed that there was no significant correlation between domain knowledge and search effectiveness.

2.4.2 Browsing Practice

Before performing the main browsing tasks, participants were asked to browse dmoz (http://dmoz.org/) and to look for information about three non-health related topics (a recipe, a football team and a music band). The purpose of this task was two-fold: first, to provide a practice browsing session for the participants and second, to observe whether participants had sufficient browsing experience. Since this was a practice task no data was recorded.

2.4.3 Main Browsing Tasks and Post-browsing Questionnaires

The browsing task consisted of 72 information search tasks (6 tasks on each of the 3 depth treatments for two different hierarchical sets of pages) on health-related topics. The number of tasks was chosen in order to provide enough power to model the browsing behavior of the participants. The experiment was designed to allow breaks at the end of every task to minimize fatigue and boredom and maintain user's interest in the browsing process [33]. In line with previous studies [13, 17] the tasks were not time-limited to enable participants to find the answer without being under time stress.

Participants were given booklets of item cards containing various health-related topics which served as targets and were asked to browse the specific hierarchy of online information presented to them and to write down the unique portion of the URL when they reached the target pages. The order of presentation was counter-balanced across participants and hierarchical structures.

After finishing the browsing tasks with each hierarchical structure, participants were asked to fill a bipolar post-browsing questionnaire on user satisfaction, orientation and the system's ease-of-use.

3 Results and Analysis

3.1 Subjective Ratings

In the main browsing experiment participants were asked to rate the expandable and non-expandable hierarchical structures in two ways: by ranking their ease of use and by rating in five bipolar scales each of the hierarchical structures in terms of ease of navigation, sense of orientation and users' satisfaction. Participants were also asked to rate their preferences in terms of the depth of the hierarchies.

3.2 Ease of Navigation

Participants were asked to rate how easy it was to navigate when browsing on a particular hierarchical structure, from 1 (very difficult) to 5 (very easy). The means and standard deviations for each age group and hierarchical structure are displayed in Table 3.

An Analysis of Variance (ANOVA) was performed on these ratings using the General Linear Model (GLM) of SPSS using a 2x3x2 (age x depth x structure) design. The results show a significant Depth main effect ($F(2,44) = 46.03$, $p<0.001$), a significant Structure*Age group effect ($F(1,45) = 5.44$, $p<0.025$) and a significant Depth*Structure*Age group effect ($F(2,44) = 3.76$, $p<0.032$). On the other hand, no significant between subject Age effect ($F(1,45) = 3.91$, $p>0.15$) was found, nor any significant Structure main effect ($F(1,45) = 2.05$, $p>0.15$) or Depth*Age group ($F(2,44) = 0.37$, $p>0.69$) or Depth*Structure ($F(2,44) = 5.44$, $p>0.44$) effects were found. Fig. 5 - Fig. 8 show the above analysis in graphical form.

Table 3. Means and standard deviations of ease of navigation

		N	Young mean (s.d.)	Old mean (s.d.)
Non-expandable	Depth2	48	4.58 (0.78)	4.21 (0.72)
	Depth 3	48	3.67 (1.13)	3.88 (0.80)
	Depth 6	48	3.04 (1.20)	3.33 (1.13)
	Overall	144	3.80 (1.22)	3.80 (0.96)
Expandable	Depth 2	48	4.63 (0.71)	4.33 (0.56)
	Depth 3	48	4.17 (1.09)	3.54 (1.02)
	Depth 6	48	3.91 (0.95)	3.21 (1.10)
	Overall	144	4.20 (0.96)	3.70 (1.03)

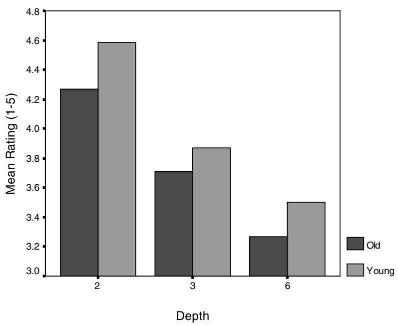

Fig. 5. Mean rating for ease of navigation by depth for the two age groups

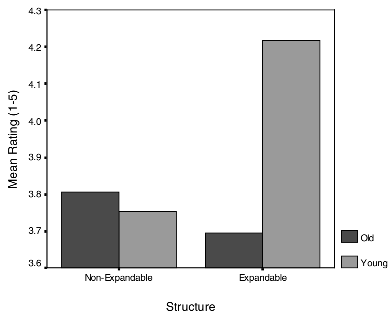

Fig. 6. Mean rating for ease of navigation for the two structures

3.3 Sense of Orientation

In this part of the experiment, participants were asked to rate their sense of orientation (in the sense of knowing where they are in the hierarchy) while browsing a certain structure, from 1 (very disoriented) to 5 (very oriented). The means and standard deviations for each age group and structure are displayed in Table 4.

An ANOVA was performed on these ratings using the General Linear Model (GLM) of SPSS using a 2x3x2 (age x depth x structure) design. The results show a significant Depth main effect ($F(2,44) = 35.76$, $p<0.001$), a significant Structure*Age group effect ($F(1,45) = 4.50$, $p<0.04$) and a significant Depth*Structure*Age group effect ($F(2,44) = 4.57$, $p<0.02$) and a significant between-subject Age effect ($F(1,45) = 4.77$, $p<0.035$) was found. On the other hand, no significant Structure main effect ($F(1,45) = 1.40$, $p>0.24$) or Depth*Age group ($F(2,44) = 0.075$, $p>0.92$) or Depth*Structure ($F(2,44) = 2.66$, $p>0.08$) effects were found. Fig. 9 - Fig. 12 show in graphical form the above analysis.

Table 4. Sense of orientation

		N	Young Mean (s.d.)	Old mean (s.d.)
Non-expandable	Depth2	48	4.71 (0.69)	4.29 (0.69)
	Depth 3	48	3.92 (1.06)	3.96 (0.86)
	Depth 6	48	3.29 (1.12)	3.21 (1.22)
	Overall	144	4.00 (1.13)	3.80 (1.04)
Expandable	Depth 2	48	4.63 (0.65)	4.33 (0.56)
	Depth 3	48	4.25 (1.07)	3.46 (1.06)
	Depth 6	48	4.17 (0.89)	3.33 (1.55)
	Overall	144	4.40 (0.90)	3.70 (1.20)

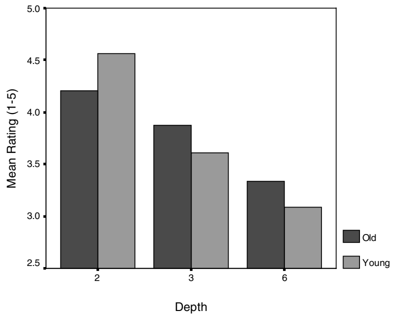

Fig. 7. Mean rating of ease of use for the non-expandable structure

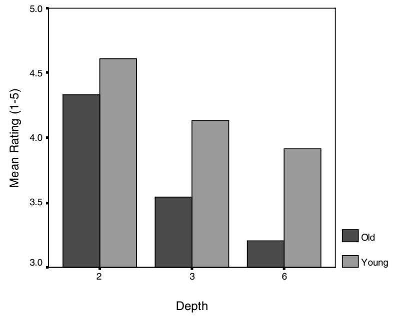

Fig. 8. Mean rating for ease of use for expandable structure

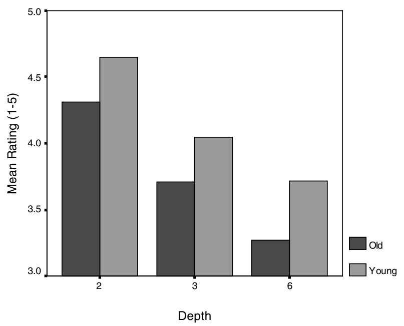

Fig. 9. Mean rating for sense of orientation by depth

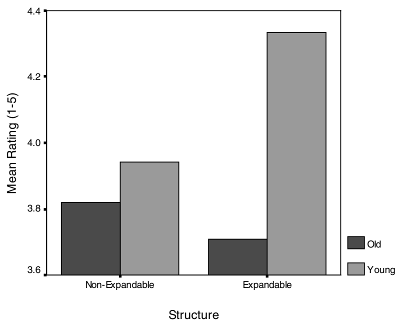

Fig. 10. Mean rating for sense of orientation for the two structures

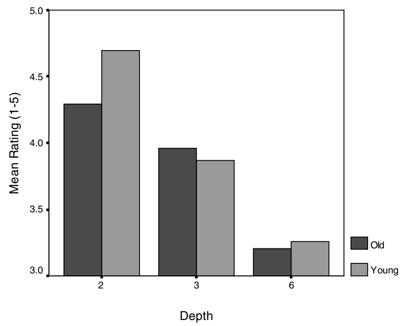

Fig. 11. Mean rating of sense of orientation for the non-expandable structure

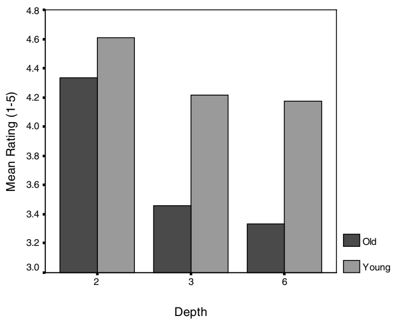

Fig. 12. Mean rating for sense of orientation for expandable structure

3.4 Satisfaction

In this measure, participants were asked about their satisfaction when using a particular structure to browse for the information, from 1 (very unsatisfied) to 5 (very satisfied). The means, standard deviations for each age group and structure are displayed in Table 5.

An Analysis of Variance (ANOVA) was performed on these ratings using the General Linear Model (GLM) of SPSS using a 2x3x2 (Age x Depth x Structure) design. The results show a significant Depth main effect ($F(2,44) = 31.86$, $p<0.001$), a significant Structure main effect ($F(1,45) = 7.44$, $p<0.01$), a significant Structure*Age group effect ($F(1,45) = 12.05$, $p<0.001$) and a significant Depth*Structure*Age group effect ($F(2,44) = 3.34$, $p<0.05$) but no significant between subject Age main effect ($F(1,45) = 2.49$, $p>0.334$) was found. Similarly no significant Depth*Age group ($F(2,44) = 0.083$, $p>0.92$) or Depth*Structure ($F(2,44) = 0.87$, $p>0.42$) effects were found.

Table 5. User satisfaction

		N	Young mean (s.d.)	Old mean (s.d.)
Non-expandable	Depth2	48	4.50 (0.88)	4.25 (0.79)
	Depth 3	48	3.79 (1.14)	3.96 (0.91)
	Depth 6	48	3.00 (1.29)	3.37 (1.10)
	Overall	144	3.80 (1.26)	3.90 (1.00)
Expandable	Depth 2	48	4.54 (0.88)	4.37 (0.65)
	Depth 3	48	4.46 (0.72)	3.75 (1.03)
	Depth 6	48	3.96 (0.98)	3.25 (1.19)
	Overall	144	4.30 (0.89)	3.80 (1.07)

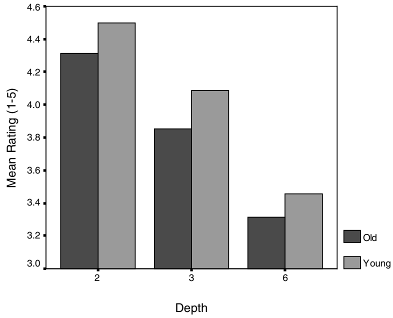

Fig. 13. Mean rating for satisfaction by depth for the two age groups

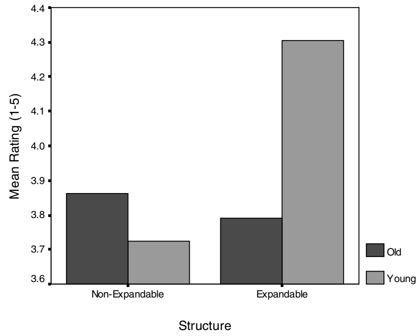

Fig. 14. Mean rating for satisfaction for the two structures

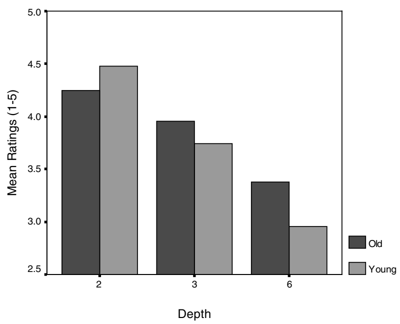

Fig. 15. Mean rating for satisfaction for the non-expandable structure

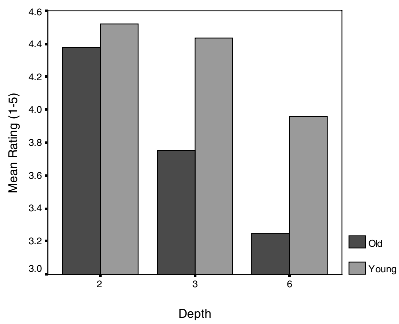

Fig. 16. Mean rating for satisfaction for expandable structure

3.5 Ranking in Ease of Use

As a measure of overall usability of the different treatments used in this experiment, participants were asked to rank those structures in terms of their ease-of-use. Sixteen younger and 14 older users ranked non-expandable structure as more difficult to use. Participants were also asked to rank their preferences in terms of depth. Both seniors and young participants preferred shallow to deep hierarchies as shown in Table 6.

3.6 Objective Measures

Browsing effectiveness (Search Time or time to complete a task) and browsing accuracy (Click Errors) are the main two objective measures of the main browsing experiment. Both of these results will be used later on to validate the different proposed models. The search time and number of clicks to target were obtained from the logs of the web servers that served the pages.

Table 6. Preference (number of participants per cell) in terms of depth

			1st Choice	**2nd Choice**	**3rd Choice**
Young	Expandable	Depth 2	15	3	3
		Depth 3	3	14	4
		Depth 6	3	4	14
	Non-expandable	Depth 2	19	3	2
		Depth 3	2	19	3
		Depth 6	3	2	19
Old	Expandable	Depth 2	11	6	6
		Depth 3	3	15	5
		Depth 6	9	2	12
	Non-expandable	Depth 2	12	3	7
		Depth 3	2	17	3
		Depth 6	8	2	12

3.6.1 Number of Clicks to Target

Table 7 lists the means and standard deviations of number of clicks to target for each combination of age group, structure and depth treatment. As expected, the ANOVA analyses show a significant main effect of Depth (F(2,45)=137.53, p<0.001). The ANOVA also shows a significant Structure main effect (F(1,46)=11.387, p<0.003). On the other hand no significant Age effect (F(1,46)=0.430, p>0.5) or Depth*Age group (F(2,45)=1.926, p>0.15), Structure*Age group (F(1,46)=1.083, p>0.3), Depth*Structure (F(2,45)=2.258, p>0.11) or Depth*Structure*Age group (F(2,45)=0.936, p>0.4) effects were found.

Table 7. Mean number of clicks to target for each treatment

Age Group	Depth	Structure	Mean (s.d.)
Young	6	Non-Expandable	11.93 (1.13)
		Expandable	9.23 (0.65)
	3	Non-Expandable	5.26 (0.39)
		Expandable	4.45 (0.18)
	2	Non-Expandable	2.74 (0.22)
		Expandable	2.45 (0.13)
Old	6	Non-Expandable	10.64 (1.13)
		Expandable	9.66 (0.65)
	3	Non-Expandable	4.60 (0.39)
		Expandable	4.04 (0.18)
	2	Non-Expandable	2.85 (0.22)
		Expandable	2.40 (0.13)

3.6.2 Search Effectiveness

Search effectiveness is described as the total time needed to arrive at the correct target. Table 8 lists the means (in seconds) and standard deviations of each age-group, structure and age group combination.

Table 8. Search Effectiveness (Time in seconds)

Age Group	Depth	Structure	Mean (s.d.)
Young	6	Non-Expandable	35.19 (3.86)
		Expandable	33.90 (3.56)
	3	Non-Expandable	23.91 (2.24)
		Expandable	23.37 (2.18)
	2	Non-Expandable	15.82 (1.80)
		Expandable	18.93 (1.79)
Old	6	Non-Expandable	45.93 (3.86)
		Expandable	55.55 (3.56)
	3	Non-Expandable	30.25 (2.24)
		Expandable	33.50 (2.18)
	2	Non-Expandable	26.12 (1.80)
		Expandable	24.18 (1.79)

The ANOVA analysis shows significant Age main between-subjects effect ($F(1,46)=13.689$, $p<0.002$), significant Depth main effect ($F(2,45)=77.280$, $p<0.001$), significant Depth*Age group ($F(2,45)=3.2$, $p<0.05$) and significant Depth*Structure*Age group effect ($F(2,45)=7.518$, $p<0.003$). On the other hand no significant Structure main effect ($F(1,46)=2.759$, $p>0.1$), Structure*Age group ($F(1,46)=1.77$, $p>0.195$) and Depth*Structure ($F(2,45)=0.96$, $p>0.39$) effects were found.

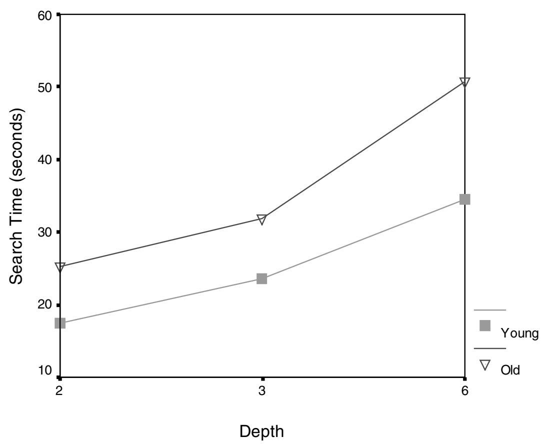

Fig. 17. Search time versus Depth for the two age groups

4 Discussions

4.1 Subjective Measures

Three measures of subjective ratings were collected in this study: ease of navigation, sense of orientation and satisfaction. The followings are the resulting observations:

When asked to rate the ease of navigating of the different treatments the following results were obtained:

1. There was a significant Depth main effect (shallow hierarchies were rated as easier to navigate than deeper hierarchies).
2. There was a significant Structure*Age group effect (older participants found non-expandable hierarchies easier to navigate whereas younger participants found expandable hierarchies easier to navigate).
3. There was a significant Depth*Structure*Age effect (younger participants rated the depth 2 expandable hierarchy as the easiest to navigate and depth 6 non-expandable as the hardest to navigate, whereas the older participants found the depth 2 non-expandable hierarchy as the easiest to navigate and the depth 6 expandable as the hardest to navigate).
4. No significant Structure main effect or Age between subject effect was found.

When the participants were asked to rate their sense of orientation (knowing where they are in the hierarchy) while browsing the different treatments the following results were obtained:

1. There was a significant between-subjects Age effect. Older participants rated their sense of orientation while performing the experiment lower than their younger counterparts.
2. There was a significant Depth main effect. Participants reported higher ratings for sense of orientation for shallow hierarchies.
3. There was a significant Structure*Age group effect. Younger participants stated that they had a better sense of orientation with expandable hierarchies whereas older participants showed a preference to non-expandable structures when it comes to sense of orientation.
4. There was a significant Depth*Structure*Age group effect. Younger participants rated the hierarchy with depth 2 of the non-expandable structure as the best in terms of sense of orientation, the older participants on the other hand rated the depth 2 structure of the expandable hierarchy as best. But both younger and older participants rated the hierarchy with depth 6 of the non-expandable structure lowest in terms of sense of orientation.
5. No significant Structure main effect, Depth*Age effect and Depth*Structure effect was reported.

When the participants were asked to rate their general satisfaction while browsing the different treatments the following results were obtained:

1. There was a significant Depth main effect. Participants reported higher general satisfaction ratings for shallow than deep hierarchies.
2. There was a significant Structure main effect. Participants reported higher general satisfaction ratings for expandable than non-expandable structures.

3. There was a significant Structure*Age group effect. Younger participants showed a preference to expandable structures whereas older participants showed a preference to non-expandable structures when it comes to general satisfaction.
4. There was a significant Depth*Structure*Age group effect. Both young and old participants rated as best in terms of satisfaction the depth 2 expandable structure. The young group rated the depth 6 non-expandable structure as the worse in terms of satisfaction, whereas the older group rated the depth 6 expandable structure as the worse in satisfaction.
5. There was no significant Depth*Age group, or Depth*Structure effect.

Finally, as a measure of overall usability of the different treatments participants were asked to rank the different hierarchies in terms of their ease of use. The results are very interesting. Both young and old participants chose the hierarchies with depth 2 as their first preferences, the hierarchies with depth 3 as their second choice, and rated the hierarchies with depth 6 as the most difficult to use.

An important observation, is that older participants in general rated all treatments much lower than their younger counterparts. For example, although there is a Structure*Age group effect for overall satisfaction (older participants preferred non-expandable hierarchies but younger participants showed a preference to expandable hierarchies) it should be noted that the rating older participants gave to both expandable and non-expandable arrangements were lower that the respective rating given to these hierarchies by the younger participants. This observation, indicates that there are further dimensions (apart from the structuring of the information) that need to be taken into consideration when presenting information online for senior citizens.

4.2 Objective Measures

Objective measures consist of two items: search effectiveness (total browsing time to arrive at the correct target) and the search efficiency (actual number of clicks needed to reach the target).

4.2.1 Number of Clicks to Target

When analyzing the data relating to the number of clicks to reach the target, obtained from the server logs, the following results were obtained:

1. There was a significant main effect of Depth. As expected, users browsing deeper hierarchies needed more steps to reach the target than did users browsing shallow hierarchies.
2. There was a significant Structure main effect. Users browsing non-expandable hierarchies took on average significantly more steps to reach the target than users browsing expandable hierarchies did.
3. No significant Age main effect, Depth*Age group, Structure*Age group, Depth*Structure or Depth*Structure*Age group effects were found.

The last result suggest that, in contrary to findings from previous studies [27], in this particular task environment older adults did not make more errors than their younger counterparts and that any age-related differences in search time are not dependent on any differences in the number of clicks to target between the two age groups.

4.2.2 Search Effectiveness

Search effectiveness was the main investigation of the study. The GLM ANOVA analysis of search effectiveness showed that:

1. There was a significant Age main effect. Seniors on average took more time to reach the target in the browsing tasks.
2. There was a significant Depth main effect. Participants on average took longer to browse deep than shallow hierarchies.
3. There was a significant Depth*Age group effect. Older participants were slower in all depth hierarchies, and disproportionately effected by increasing depth.
4. There was a significant Depth*Structure*Age group effect. Younger participants browsed the depth 2 non-expandable hierarchy the fastest and the depth 6 non-expandable hierarchy the slowest. On the other hand the older participants browsed the depth 2 expandable hierarchy the fastest and the depth 6 expandable hierarchy the slowest.
5. There was no significant Structure main effect, Structure*Age group effect and Depth*Structure effect.

The aforementioned results suggest that information designers should take into consideration the fact that older users take longer time compared to younger users to find the answers they were looking for in hierarchical information structures. This result shows the importance of information architecture in optimizing the user experience online.

The results also show that, in contrary to findings from previous studies (e.g. [13]), users do not perform significantly better in the expandable hierarchy. Unlike the Zaphiris, Shneiderman and Norman [13] experiment; all the stimuli in the experiment described in this paper could fit into a single screen when fully expanded (in the case of the expandable hierarchies). This avoided the trouble of scrolling which the Zaphiris, Shneiderman and Norman's [13] study found to be the main reason of the difference in search effectiveness between expandable and non-expandable hierarchies.

5 Conclusions

The focus of this study was hierarchical online information structures. Previous studies have shown that investigating such structures and the depth versus breadth tradeoff in browsing hierarchical structures can provide valuable information to designers and researchers to better design online information. Information architecture (the art/science of optimally designing/arranging information online) is a crucial element in this task.

The results show that there was a significant Depth main effect with participants taking longer to reach the target in deep than shallow hierarchies. Also there was a significant Depth*Age group effect, with older participant being slower in all depth treatments. Furthermore, Shallow hierarchies were preferred to deeper hierarchies for ease of navigation, sense of orientation, and overall satisfaction.

Finally, a significant Structure effect was reported for general satisfaction with expandable hierarchies rated higher than non-expandable structures and a significant Structure*Age group effect reported for ease of navigation, sense of orientation, and general satisfaction ratings. Older participants preferred the non-expandable hierarchies, whereas their younger counterparts preferred the expandable hierarchies.

Overall, seniors were slower and did make more mistakes than their younger counterparts when browsing the different treatments.

5.1 Suggestions to Researchers and Practitioners

There are several lessons learned from this study that are fruitful contributions to the area of human-computer interaction, especially for older information users. First of all, as reported in the above analysis, there are age-related differences both in terms of preference and in terms of performance when it comes to browsing hierarchical online information structures. This result should be taken seriously by web designers, especially when designing online information that targets senior citizens. Where possible, shallow hierarchies should be preferred to deeper hierarchies and when targeting seniors, non-expandable hierarchies should be preferred.

5.2 Limitations

First (and most importantly) the current study was concentrated on homogeneous structures. Future research could focus on strengthening these results by running similar studies for more complex (non-homogeneous) structures. Secondly, this study was limited in that it was tested under the special case of text-only websites. Web sites usually contain banner ads, pictures and icons. The effect of such visual elements was not investigated in this study. Finally, there is obviously a necessity for accurate mathematical models that could be used with confidence in predicting the complexity and thus the number of clicks to target when browsing hierarchical structures.

References

1. Rosenfeld, L., *Special Report: Design Usability – Seven Pitfalls to Avoid in Information Architecture*. 2000.
2. Wurman, R.S., *Information Architects*. 1996, Zurich: Graphics Prss.
3. Davenport, T.H., *Information Ecology: Mastering the Information and Knowledge Environment*. 1997, New York: Oxford University Press.
4. Korman, R., *Helping users find their way by marking your site 'smelly'*. 1998, webreview.com.
5. Nielsen, J., *User interface directions for the web.* Communications of the ACM, 1999. **42**(1): p. 65–72.
6. Gullikson, S., et al., *The impact of information architecture on academic web site usability.* The Electronic Library, 1999. **17**(5): p. 293–304.
7. Elm, W. and D. Woods. *Getting lost: A case study in interface design.* in *The Human Factors Society 29th Annual Meeting*. 1985.
8. Kim, H. and S.C. Hirtle, *Spatial metaphors and disorientation in hypertext browsing.* Behaviour and Information Technology, 1995(14): p. 239–250.

9. Edwards, D. and L. Hardman, *'Lost in Hyperspace': Cognitive mapping and navigation in a hypertext environment*, in *Hypertext: Theory into Practice*, R. McAleese, Editor. 1989, Intellect: Oxford. p. 90–105.
10. Simpson, A. and C. McKnight, *Navigation in hypertext: structural cues and mental maps*, in *Hypertext: State of the Art*, G. C., Editor. 1990, Intellect: Oxford. p. 73–83.
11. McDonald, S. and R.J. Stevenson, *Disorientation in hypertext: The effects of three text structures on navigation performance.* Applied Ergonomics, 1996(27): p. 61–68.
12. Fleming, J., *Web Navigation: Designing the User Experience*. 1998, Sebastopol: O'Reilly.
13. Zaphiris, P., B. Shneiderman, and K. Norman, *Expandable versus sequential menus.* Behaviour and Information Technology, In Press.
14. Snowberry, K., S.R. Parkinson, and N. Sisson, *Computer Display Menus.* Ergonomics, 1983. **26**(7): p. 699–712.
15. Larson, K. and M. Czerwinski. *Page Design: Implications of Memory, Structure and Scent for Information Retrieval*. in *CHI '98 Human Factors in Computing Systems*. 1998: ACM Press.
16. Zaphiris, P. *Depth vs Breadth in the Arrangement of Web Links*. in *Human Factors Society, Fourty-Fourth Annual Meeting*. 2000. San Diego.
17. Wallace, D., N. Anderson, and B. Shneiderman. *Time Stress Effects on Two Menu Selection Systems*. in *Human Factors Society, Thirty-First Annual Meeting*. 1987.
18. Jacko, J. and G. Salvendy, *Hierarchical Menu Design: Breadth, Depth, and Task Complexity.* Perceptual and Motor Skills, 1996(82): p. 1187–1201.
19. Miller, D.P. *The Depth/Breadth Tradeoff in Hierarchical Computer Menus*. in *Human Factors Society*. 1981.
20. Kiger, J.I., *The Depth/Breadth Tradeoff in the Design of Menu-Driven Interfaces.* International Journal of Man-Machine Studies, 1984(20): p. 201–213.
21. Parkinson, S.R., N. Sisson, and K. Snowberry, *Organization of broad computer menu displays.* International Journal of Man-Machine Studies, 1985. **23**: p. 689–697.
22. McDonald, J., J. Stone, and L. Liebelt. *Searching for items in menus: The effect of organization and type of target*. in *Human Factors Society, 31st Annual Meeting*. 1983.
23. Nygren, E., *"Between the clicks" Skilled Users Scanning of Pages*. 1996.
24. Norman, K. and J. Chin, *The effect of tree structures on search in a hierarchical menu selection system.* Behaviour and Information Technology, 1988. **7**: p. 51–65.
25. Campbell, D.J., *Task Complexity: a review and analysis.* Academy of Management Review, 1988(13): p. 40–52.
26. Paap, K.R. and R.J. Roske-Hofstrand, *The Optimal Number of Menu Options per Panel.* Human Factors, 1986. **28**.
27. Mead, S., et al. *Effects of Age and Training on World Wide Web Navigation Strategies*. in *Human Factors Society, Forty-First Annual Meeting*. 1997.
28. Administration on Aging, *A Profile of Older Americans: 2001*. 2002.
29. White, H., et al., *Surfing the net in Later Life: A Review of the Literature and Pilot Study of Computer Use and Quality of Life.* The Journal of Applied Gerontology, 1999. **18**: p. 358–378.
30. Cochrane, J.D., *Healthcare @ the Speed of Thought.* Integrated Healthcare Reports, 1999: p. 16–17.
31. Zaphiris, P. and R.D. Ellis. *Mathematical Modeling of Age Differences in Hierarchical Information Systems*. in *ACM Conference on Universal Usability*. 2000. Arlington.
32. Zaphiris, P. and S.H. Kurniawan. *Using Card Sorting Technique to Define the Best Web Hierarchy for Seniors*. in *CHI 2001 Conference on Human Factors in Computing Systems*. 2001. Seattle.
33. Hirashima, T., et al. *Context-Sensitive Filtering for Browsing in Hypertext*. in *International Conference on Intelligent User Interfaces*. 1998. New York: ACM Press.