Learning a menu selection tree: training methods compared

DIANA PARTON‡, KEITH HUFFMAN‡, PATTY PRIDGEN‡,
KENT NORMAN‡ and BEN SHNEIDERMAN‡§

†Department of Computer Science and ‡Department of Psychology,
University of Maryland, College Park, MD 20742, U.S.A.

Abstract. Menu selection systems sometimes present learning problems for novice
users. This comparison of four training methods for novice users found that the
global tree diagram of the menu system was superior to command sequence and
frame presentation methods, and somewhat better than trial and error. Methods
were evaluated on the basis of (1) number of target nodes found, (2) mean number of
selections to a target node, (3) recall of the menu structure, and (4) subjective rating
of ease of learning.

1. Introduction
Menu selection systems have become increasingly popular as a means of
information retrieval, particularly for inexperienced computer users (Shneiderman
1985). One advantage of this method is technological: menu selection makes the most
efficient use of computer resources for the display of large quantities of information
(McEwen 1982). In a hierarchical system, very general categories of information are
successively broken down into increasingly specific subcategories until the desired
frame of information is reached. Users are therefore not required to memorize specific
command sequences for information access. Liebelt et al. state that the ‘user must only
recognize the correct option rather than recall the appropriate command’ (Liebelt et al.
1982).

1.1. Learning problems of menu selection systems
Advocates of menu selection systems often focus on the ease with which novice
users are believed to be able to learn and use them. However, most novice users
experience some degree of difficulty in interacting with a menu selection system.
Sometimes the difficulty is due to the user’s method of searching while studying a
particular frame. Card (1982) found that most intra-frame searching is done randomly,
not by systematic up-and-down eye movements. The user scans the frame by blocks he
labelled ‘perceptual chunks’. These chunks require a variable number of eye
movements. Depending on the frame structure there is a difference in visual search
times. Card determined that frame architecture was a significant factor in search time
within a single menu. He concluded that alphabetic frame design permitted faster
searches when compared with functional frame design. Both designs proved to be
significantly faster means than random frame design.

The organization of menu frames is a major factor in the ease of menu system use.
Liebelt et al. (1982) determined that organized frames in a menu system influenced ease
of learning. This is supported by Bower et al. (1969) who confirmed that more items are
recalled from organized structures than from random ‘permutations’ of the structures.
Bower’s data were not as conclusive as Liebelt’s solely for organization since there were

§ Author for correspondence.
meaningful 'category names' resulting in close associations to 'category examples' (Liebelt et al. 1982). From the Liebelt and Card results, Liebelt stated, 'It would appear, therefore, that the combination of within- and between-panel (frame) organization would yield an optimal menu structure'.

Another problem faced by the novice user is the difference in perception of the menu system by the user from that of the designer (Billingsley 1982, Dumais and Landauer 1982). Young (1982) has termed this difference 'cognitive mismatch'. The difficulty of responding to the menus may be attributed to any of the following: the physical layout is misleading relative to the context intended, the classification structure is obscure, or the users react before carefully reading the frame. Misconceptions can result in users becoming disoriented in the menu system. Initial user confusion can lead to persistent trouble throughout the search. The confused user seems to require, on the average, twice the minimum number of frames necessary to locate a terminal target (McEwen 1982).

1.2. Proposals to make learning easier

Many solutions for such problems have been proposed. Dray et al. (1981) believe that the traditional menu–calling–submenus approach is difficult for novices to follow. He recommends the use of multiple-item line menus, particularly in cases where accuracy is important to the search and speed of retrieval is not a factor. When the user is comfortable with the line menu, Dray recommends that the user is then ready to learn to use the hierarchical menu form.

The question of menu 'shape' is also relevant to user disorientation. Miller (1981) experimented with menu configurations of assorted depth (number of levels to the information frames) and breadth (number of choices within a particular menu frame). In his menu system with 64 terminal nodes, the configuration that yielded fewest errors and fastest information retrieval time was a two-level system where each frame presented eight choices. Miller recommends a broad and shallow design as superior to those of increased depth. He believes it is difficult to retain path lengths greater than five frames long in users' short-term memory. The number of levels in a menu system should be minimized, but not at the expense of cluttering each frame display with too many choices.

Numerous searching aids available to the user during the retrieval process have been devised and tested. McEwen (1982) provided menu users with either a Name-Trace feature which listed up to five frame titles of previously selected frames above the currently displayed frame, or a Previous Page feature displaying the entire preceding page above the current frame. Results for both these groups, as well as a control group with access to neither feature, were very similar in terms of mean search time, mean number of key-strokes made, and mean number of page accesses. However, the Name-Trace group made significantly fewer first response errors in retrievals involving returning two or more levels back up the tree, while Previous Page subjects made significantly fewer first response errors in search strategies requiring retracing only one level.

As an alternative to the hierarchical tree-search technique, Tombaugh and McEwen (1982) provided an alphabetic directory of the data base being searched. Performance for groups using one or the other retrieval strategy—or a combination of both—was not significantly different, although there was a strong tendency to switch from the tree-search strategy into the alphabetic directory more often than vice versa.
Major public information retrieval networks such as The Source and CompuServe provide such alternatives to their subscribers. Desired information can be located in either network by interactive menu selection or by a command which directly accesses the desired page of information. It is possible to switch from menu to commands or vice versa at any time.

Magid (1983) in his comparison of CompuServe and The Source points out that menus are recommended for new service users, since they provide more information and are easier to learn than commands. The Source, which was originally command-driven, added menus to supplement the command display (Derfler 1982). The disadvantage of menu selection, particularly when subscribers pay by the hour for the information service, is the relatively slow menu display. For subscribers who prefer more direct access, The Source provides a Directory of Services and CompuServe supplies an Index of current command listings. These are also available to subscribers in printed form through new user documentation and update mailings. CompuServe's Index is available on-line directly from the Main Menu.

Electronic Information Exchange System (EIES) is another large information network 'designed to study and explore the use of computers to facilitate human communications across a wide variety of applications' (Hiltz and Turoff 1981). These communications components include public, specific group, and private conferences, messages, and notebooks. The system design was 'intended to allow user groups to evolve features tailored to the nature of a group and its application'. Users can access information in one of four modes: Menu Selection, taught to users in the system documentation; Commands, where all menu selections as well as 200 other 'advanced features' are available, Answer-Ahead and Command Streams, enabling the user to commence a series of operations or anticipate future questions and answers; and Self-Defined Commands, where an individual or group can specify a unique command.

Hiltz and Turoff conducted a two-year study of changes in user behaviour and usage patterns of scientists using EIES voluntarily in the course of their work. Hiltz and Turoff claimed that EIES did not provide adequate documentation for command features and that advanced features were sought out only by those strongly motivated to do so. 'EIES members feel that the menu is the optimal interface for the beginning user'. However, with increased experience, users most valued 'features that allow a user to actively control the system rather than to passively react to menu choices' when they understood the options available to them.

McGee (1976) proposed that a pictorial representation could help with initial learning of data models, and thereafter remain as a reference for users (Billingsley 1982). Fitter and Green (1979) believed that pictures were faster and easier to assimilate than purely semantic representations.

Billingsley built on this base when she provided 'maps' through a data base to the target entries. She believed that users who were able to study the menu structure as the system designer envisioned it would be able to use the system correctly and remember menu sequences more easily. 'Problems created by ambiguity in menu choices are compounded when users are unaware of the hierarchical structure of data organization'. She provided one group of subjects with a Data Index, 'specific semantic pathways' through the hierarchy to target words, by listing each target word and the entire sequence of menu choices required to reach it. A second group studied a Data Map of the hierarchy showing all the pathways and relationships between the data. A control group with no such information made significantly more choices to reach target words than either the Map or Index groups. Index and Map groups both performed
well with the initial group of targets to be found. However, after a new group of target words was introduced, the Map group was significantly faster in reaching the desired target words. Whereas the Index group and particularly the control group had to rely on semantic relationships only, Billingsley believed the maps added mnemonic information that the users retained even after the physical maps were no longer available, and this accounted for the difference in performance.

1.3. *An experiment in training methods*

The purpose of the present experiment was to try to determine the best method of presenting the information contained in a menu system to aid users learning the system. Four different sets of materials were prepared. Each set represented, in different formats, the information contained in a hierarchical menu system. The subjects’ task was to study their materials and then use the same menu selection system to find program-supplied target words at the lowest level of the hierarchy.

The method of presentation of the study material was varied across the following conditions:

1. The Trial and Error group was given no printed study material; instead, they were allowed to use the study period to go through a trial run of the menu selection system. They simply explored to see where different submenus would take them.
2. The Command Sequence group was given a one-page list of the pathways that led through the hierarchy to the target words.
3. The Frame Presentation group was given one sheet of study materials which contained all the frames from the menu system in random order.
4. The Global Tree group was given a diagram representing the overall structure of the menu system, with all frames connected in tree fashion.

Four dependent variables were assessed to measure the effectiveness of each method of menu presentation:

1. number of target words found,
2. number of targets recalled in a post-experiment reconstruction task,
3. average number of tries to reach target words, and
4. user assessment of ease of learning.

The number of target words found and the number recalled were used to measure the ease in learning and remembering a particular method. Likewise, the number tried provided a measure of how clear the method of presentation was. Finally the subjective measure of user satisfaction was used to assess user impression of how easy it was to learn and use the system.

2. *Experimental design and procedure*

2.1. *Subjects*

The subjects consisted of 65 University of Maryland undergraduate psychology students who participated in the experiment for credit in their psychology classes. Subjects were told that computer experience was not necessary. Most subjects had no such previous experience. One subject’s set of data was discarded because examination of his listing revealed that he was deliberately entering invalid choices.

The menu selection system itself was implemented as a ternary tree, three levels deep. It represented a fictitious organizational hierarchy with 'divisions' as Main Menu
selections, 'sections' on the second level, and specific job titles at the lowest level. Menu entries were in reasonably logical relationships, but study was necessary to learn the precise sequences from the Main Menu to the target titles. It was intended that subjects would not be able to rely on strictly logical relationships to find their way through the system.

Two computer programs were written to allow the subjects to traverse the menu structure. Both programs printed the menu frame title and the choices available at that frame. All frames had three menu choices, and every frame except the Main Menu had a fourth choice which allowed the user to return to the previous frame.

One program supplied the target word the subjects were to search for. In case the subjects forgot the target word, a fifth choice of entering 't' to cause the target word to be repeated was allowed. The subjects were told that this choice was available, but it did not appear as an option in the menu frames presented. When a correct target word was found, the program acknowledged the user's success, and then returned the user to the Main Menu frame where a new target word was supplied. An advisory message was printed when an incorrect target word was reached. All subjects took the experiment using this program.

A modified version of this program was implemented to permit the Trial and Error group to move freely through the menu system during their study period. No target words were presented to subjects in this program. When the lowest level of the tree structure was reached, a message to that effect was printed to prompt subjects to return to the next higher level.

Each subject was given a folder with written instructions describing the experimental procedures. Everyone had identical general instructions concerning the purpose of the experiment. In addition, the instructions explained how to make a selection and enter it into the computer, and presented an interactive example of traversing the menu structure starting at the main menu and finishing at the lowest level.

Subjects were also given written instructions and study materials specific to the presentation method they would be using. Each subject was assigned to one of the following presentation method groups:

*Trial and Error.* This group was not given any study material. Instead, subjects were allowed to go through the on-line menu system for the 12-minute study period.

*Command Sequence.* Subjects in this group studied a one-sheet listing of the job title sequences that progressed from the Main Menu to the lowest level of the hierarchy. The sequences were arranged in random order.

For example:

Plants Division, Concepts, Systems Analyst.

*Frame Presentation.* This group was given the 13 frames, randomly ordered, that comprised the tree structure used in the menu system. All frames were on the same page. Each frame had a box drawn around it to delimit it from the others, and each frame's name was underlined. For example:

<table>
<thead>
<tr>
<th>Plans Division</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concepts</td>
</tr>
<tr>
<td>Designs</td>
</tr>
<tr>
<td>Proposals</td>
</tr>
</tbody>
</table>
**Global Tree.** These subjects studied a diagram of the tree structure representing the hierarchy implemented by the system. Each 'node' in the tree was a menu frame at the appropriate tree level. Each frame had a box drawn around it, and each frame's name was underlined. Straight lines connected each parent node with its children.

Following the test, two additional sets of materials were used. Subjects were given a rating scale from 1 (very difficult) to 7 (very easy) to indicate the ease of learning the menu system based on the method of presentation.

Secondly, subjects were given a tree diagram that represented the menu system. Each node in the tree was drawn as a box with space for the name of the frame and its three choices. The Main Menu was filled in for the user, and its three entries were then put in the name position of their respective frames. This was done to give the subjects an idea of the ordering of the frames and the representation of the menu system as a tree structure.
2.3. Administration

The experiment was conducted on Digital DecWriter II terminals on a large mainframe computer. These produced readable dot-matrix output on paper at 30 characters per second. The Dec-Writers were covered to limit the subjects' view to only the current frame. The window was approximately four inches high. Subjects participated in groups of six to twelve students, during a time when user demand on the computer was fairly equal and system response did not vary too much across the experiment sessions. In addition, subjects in the Command Sequence, Frame Presentation and Global Tree conditions were run at the same time so as to match subjects on system response. These conditions could be mixed since the study period for these methods required memorization of the respective lists, charts, and diagrams. On the other hand, the Trial and Error method needed no paper study materials; consequently, this condition had to be controlled separately.

To be consistent, the same experimenter read the general instructions aloud to each set of participants while they read silently. An opportunity to ask questions followed. When all questions had been answered, each subject silently read a specific set of instructions pertinent to his or her method of presentation, and began the 12-minute study period. The Trial and Error group spent their 12 minutes in an interactive exploration of the Menu Selection System, while the other three groups were instructed to memorize the elements of the Menu Selection System as it was presented to them.

After the study period, the subjects were told to close their instruction folders and begin the actual on-line accessing of the menu system. Subjects were given 10 minutes on the terminal to find as many target words as they could. They were not told how many target words were possible. At the end of the 10 minutes, they were instructed to stop and told how to exit the program.

The tree reconstruction diagrams were distributed and subjects were encouraged to fill out as many titles as possible. The rating scale for ease of learning was also passed out at this time. No time limit was imposed for this portion of the experiment.

3. Results

The mean number of targets found by subjects in each training method is listed in table 1. Although the Global Tree and Trial and Error groups found more targets than the Command and Frame groups, the difference was not statistically significant \( F(3, 61) = 1.347, \ P = 0.268 \).

The mean number of menu items recalled is listed in table 2. Recall was scored so that the subject received credit as long as the job titles were grouped together correctly at the proper level of the tree. Recall was highest in the Global Tree condition, while the other three conditions were very similar \( F(3, 61) = 3.522, \ P < 0.05 \). The Global Tree group tended to recall around 40 to 50% more items than the other groups.

The mean number of selections made until the target item was found is listed in table 3 for each training method. Although the Command Sequence group and the Frame group made nearly twice as many wrong selections until they located the target as did the Trial and Error and the Global Tree groups, the difference only approached significance \( F(3, 61) = 2.298, \ 0.05 < P < 0.10 \).

Finally, the mean ratings of ease of learning are listed in table 4 for each group. Subjects in the Global Tree group gave the average the highest ratings for ease of learning, followed by the Trial and Error group and the Command Sequence group. The lowest mean rating was given by the Frame group \( F(3, 61) = 4.263, \ P < 0.01 \).
Table 1. Mean number of targets found for each training method.

<table>
<thead>
<tr>
<th>Training method</th>
<th>Mean targets found</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>8.19</td>
<td>6.21</td>
</tr>
<tr>
<td>Command</td>
<td>4.71</td>
<td>4.18</td>
</tr>
<tr>
<td>Frame</td>
<td>6.47</td>
<td>6.26</td>
</tr>
<tr>
<td>Tree</td>
<td>8.47</td>
<td>5.51</td>
</tr>
</tbody>
</table>

Table 2. Mean number of menu items recalled for each training method.

<table>
<thead>
<tr>
<th>Training method</th>
<th>Mean recalled</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>10.13</td>
<td>8.46</td>
</tr>
<tr>
<td>Command</td>
<td>8.43</td>
<td>4.97</td>
</tr>
<tr>
<td>Frame</td>
<td>9.82</td>
<td>8.76</td>
</tr>
<tr>
<td>Tree</td>
<td>16.65</td>
<td>7.48</td>
</tr>
</tbody>
</table>

Table 3. Mean number of selections to target for each training method.

<table>
<thead>
<tr>
<th>Training method</th>
<th>Mean selections</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>10.63</td>
<td>6.39</td>
</tr>
<tr>
<td>Command</td>
<td>20.36</td>
<td>18.81</td>
</tr>
<tr>
<td>Frame</td>
<td>19.59</td>
<td>20.82</td>
</tr>
<tr>
<td>Tree</td>
<td>9.35</td>
<td>7.62</td>
</tr>
</tbody>
</table>

Table 4. Mean ease of learning rating for each training method.

<table>
<thead>
<tr>
<th>Training method</th>
<th>Mean ease of learning rating</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>3.63</td>
<td>1.58</td>
</tr>
<tr>
<td>Command</td>
<td>3.14</td>
<td>1.55</td>
</tr>
<tr>
<td>Frame</td>
<td>2.76</td>
<td>1.73</td>
</tr>
<tr>
<td>Tree</td>
<td>4.76</td>
<td>1.77</td>
</tr>
</tbody>
</table>
The correlations between pairs of dependent variables were computed (table 5). These correlations indicate that there is a fairly strong relationship among all the dependent variables. Number found and average number of selections are both performance variables and correlate highly but, as one would expect, a negative relationship exists between them. Number of items recalled correlated highly with the number found, indicating a strong relationship between performance and memory of the items in the tree. Finally, ease of learning was correlated both with the performance variables and with recall, indicating that subjects were able to assess their own performance with a certain degree of accuracy.

4. Discussion

This study was designed to evaluate four methods of presenting a menu selection system to the novice user. It is assumed that individual differences in memorization skills and strategies play a major role in the performance of the user. However, it was expected that the method of presentation would lead to significant differences in performance with the menu system, memory of the menu tree, and subjective assessments of the ease of learning the system.

Analysis indicated that recall and ratings of ease showed significant differences with respect to number of items found and the average number of selections to the target showed interesting trends but were not statistically significant by themselves. For number found the Trial and Error and Global Tree groups out-performed the Command Sequence and Frame Presentation groups. Similarly, for average number of selections the Trial and Error group and the Global Tree group out-performed the Command and Frame groups in that the first two groups required fewer selections. Taken together it seems safe to conclude that the Trial and Error and the Global Tree presentations were superior. The Global tree subject, on the average, found 8.47 target words using the lowest average tries at 9.35. Furthermore, this user tended to remember more entries in the menu tree suggesting better comprehension by subjects in the Global Tree group. It is interesting that these novice users were also aware of this advantage in that they tended to rate ease of learning highest among all four groups.

The Trial and Error results proved surprising. The users located 8.19 target words with an average of 10.63 tries. The recalled entries were somewhat less than the Global Tree group. This may be because users did not traverse the entire tree and acquire information on all of its branches. Furthermore, users tended to rate the Trial and

<table>
<thead>
<tr>
<th>Number</th>
<th>Number of selections</th>
<th>Ease of learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number found</td>
<td>$0.77$</td>
<td>$0.62$</td>
</tr>
<tr>
<td>$P &lt; 0.01$</td>
<td>$P &lt; 0.01$</td>
<td>$P &lt; 0.01$</td>
</tr>
<tr>
<td>Number recalled</td>
<td>$-0.48$</td>
<td>$P &lt; 0.01$</td>
</tr>
<tr>
<td>$P &lt; 0.01$</td>
<td>$P &lt; 0.01$</td>
<td>$-0.25$</td>
</tr>
<tr>
<td>Number of selections</td>
<td>$P = 0.02$</td>
<td></td>
</tr>
</tbody>
</table>
Error presentation as fairly easy to learn. Contrary to the hypothesis that this method would present the least desired results, such a positive showing seems quite puzzling. One possible explanation is that subjects in this group performed well because they had hands-on experience using the menu system and DecWriter terminals prior to the actual target word search, while the other three groups did not. This experience may have enabled the users to overcome some of their initial unfamiliarity with the computer. One of the major obstacles confronted by the other three methods was that subjects did not remember to press the RETURN key to enter their selection into the computer, although the instructions informed them of this and an interactive example was given. Some subjects would press the number key corresponding to the menu selection, then sit and wait, sometimes for several minutes, for the computer to respond. After a few minutes, they would ask what was wrong. They had simply forgotten to press RETURN. By the time the Trial and Error group got to the actual experiment, they knew they should do this. In addition, the Trial and Error group was familiar with how to go back to previous levels in the menu system so they were not afraid of taking wrong turns. Again, this is knowledge that users of the other three methods had to acquire during the actual execution session at the terminal.

As anticipated, the Frame Presentation and Command Sequence yielded poorer results than the Global Tree group. The Frame Presentation subjects either grasped the connection between frames or were extremely confused. They found an average of 6.47 target words requiring, on the average, 19.59 tries. The users' recall was similar to the Trial and Error group at 9.82 entries. The most striking result was that the users seem basically dissatisfied with this method of presentation, as seen by the lowest mean rating of ease of learning.

It was surprising that the Command Sequence group exhibited the least productive results. The users found only 4.71 target words using, on the average, 20.36 tries. The number of entries recalled was 8.43 suggesting confusion about the items in the tree. They also exhibited a low mean rating for ease of learning. Several factors may have contributed to this poor performance. First, the 27 command sequences were double-spaced on a single page. It is possible that the sequences were placed too close together on the sheet, and were confusing to the subjects. Secondly, the sequences were randomly ordered on the page, thus requiring more effort to discover the structure of the menu tree.

5. Conclusions
The best learning method for users would be the method which resulted in the greatest number of target words found. Likewise, the best method would result in the greatest number of correct reconstruction diagram entries. The method which required the fewest tries to reach a target word would reflect the clearest presentation. Similarly, the most easily learned method should be apparent to the user and rated as the easiest to learn.

Although Command Sequences are often preferred after users become familiar with a system, our novice users did not use them efficiently and gave them low satisfaction ratings. Presentation of frames alone, with no links to other frames, may be clear to some new users, but others cannot perceive the relationships between the frames. Trial and Error, for small systems, appears to be a good way for novices to learn the system. They gain familiarity with the system as they proceed, and learn quickly that they can recover from mistakes. The tree diagram of the entire menu structure, when practicable
in small menu systems, provides the clearest picture of the system and highest satisfaction rating for new users.

A number of issues remain to be investigated in the assessment of a superior training method. The ease of learning with the various methods may be influenced by changes in the size and shape of the menu system. Increased experience on the user’s part may make Command Sequence of frame presentation more appealing. A pre-experiment session at the terminal using a different menu system might affect the performance and preferences of novice users. It would also be interesting to determine the study strategies employed by the subjects as they studied the menu selection system.

Acknowledgments

We thank the Computer Science Center for providing document preparation computer support, Control Data Corporation for partial support of this research, and Bob Orr for technical programming assistance.

References


