The Effect of Method of Instruction and Spatial Visualization Ability on the Subsequent Navigation of a Hierarchical Data Base

Scott A. Butler
Department of Psychology
and the
Human/Computer Interaction Laboratory
University of Maryland
College Park, MD 20742

Abstract
Users need instructions about the structure of a data base and about what commands are available for use in menu traversal. Early research has shown that subjects who received tree diagrams of a data base were most successful in searching a hierarchical data base. Later research demonstrated that subjects who studied information presented in the form of an analogy performed better than did subjects who studied a tree diagram. Considering work by Vicente, Hayes, and Williges, the present experiment attempted to show that spatial visualization ability can be used to determine what form of instruction results in a difference in performance very early in the use of a data base. Subjects were grouped according to their spatial visualization ability and were randomly assigned to one of four study conditions. Subjects studied either a tree diagram of the data base of college courses, an analogy describing the structure of the data base, both the analogy and the tree diagram, or nothing about the data base. Results replicated the finding that studying a tree diagram results in the best performance among the four methods of study. Furthermore, high spatial visualization subjects made fewer moves and took less time to find targets than did low spatial visualization subjects. However, results provided no support for accelerated learning when subjects were provided with materials that complemented their spatial visualization ability.
The Effect of Method of Instruction and Spatial Visualization Ability on the Subsequent Navigation of a Hierarchical Data Base

With the growing availability of computers, more novice users are being required to interact with computers as part of their day to day activity. One of the most common modes of interaction with computers is menu oriented (Norman, in press). Tax hot-lines present callers with series of options about what information can be accessed by pressing numbers on a touch-tone phone. Automated teller machines require users to make selections in a variety of windows before their transaction is processed. Cosmetic counters are now equipped with small computers that will choose one's most complementing colors after one enters information about complexion, hair color, and eye color. All of these commonplace occurrences are examples of menu driven computer programs.

Unfortunately, not all menu systems are easy to use. As a data base becomes more complex, items must be arranged into meaningfully organized groups. The most common semantically organized design is the hierarchical design. By using such an intrinsically pleasing design, programmers create a "sensible, comprehensible, memorable, and convenient semantic organization [that is] relevant to the user's task" (Shneiderman, 1987). However, in such a design, items may belong to more than one category and the potential that the programmer's categorization matches the user's categorization of an item probably decreases as the size of the data base increases, thus increasing the potential for confusion.

Categorization differences are among only a few of the considerations a software designer must face when creating a piece of software. Once the categories have been created, the problem of how to arrange them within the hierarchy must be solved; this is the issue of depth versus breadth. Deep menus, or menus with very specific categories that are arranged in a large number of levels, are considered to be useful because they serve to funnel the users towards their target and insulate them from other information contained in the data base. Although deep menus allow for very specific semantic categorization, designing software with shallow menus, or menus with only a few levels, is seen as being the better way to present information. By presenting fewer nodes with more choices per node, the user is better able to infer the rules by which the system was organized and is able to compare and eliminate similar items (Bury, Krysiak, and Arnold, 1988). Wallace, Anderson, and Shneiderman (1987) found that reducing a hierarchical menu structure from six to three levels led to a more than six-fold decrease in errors. Seppala and Salvendy (1985) found that any hierarchical structure results in longer search time and higher error rates than parallel structures. It seems that hierarchical menu structures are difficult to navigate but, regardless of their complexity, hierarchical menus are also very common,
intrinsically appealing, and useful for organizing and presenting large amounts of data. Therefore, attention must be paid to the problem of how to instruct users to use a hierarchical data structure.

Methods of Instruction

Learning a new skill is challenging and anyone using a data base for the first time will require some sort of training. Learners typically get some information from more experienced users, but documentation, either on-line or on paper, is usually necessary for a complete understanding of the system (Shneiderman, 1987). Research investigating the topic of method of instruction has yielded a variety of results. Norman and Schwartz (1987) presented subjects with one of four types of information before having them search a content-free menu. The four types of information were a diagram of the menu tree, a list of command sequences required to arrive at desired targets, the sequences of menu screens that must be traversed to arrive at targets, and no information about the menu. No differences between mode of instruction were found for the number of targets that subjects located, but differences in the type of information used to locate a target were found. Parton, et al (1985) compared global tree, command sequence, frame presentation, and trial-and-error instructional methods. Subjects in the global tree instructional condition gave the highest "Ease of Learning" ratings for learning to use the menu system. Subjects who studied a global tree of the menu system also recalled more items from the tree in a test following the experimental searches.

Billingsley (1982) had subjects study a tree diagram, command sequences, or no information about a menu with meaningful contents. Subjects performed four blocks of 10 searches and only after subjects had completed 30 searches did significant differences appear. In the fourth block of searches, subjects who had studied the tree diagram found targets faster and had better mental models of the menu structure than did subjects who had studied other instructional material.

Later research by Webb and Kramer (1988) examined the performance of subjects who were provided with one of the following instructional aids: no information about the menu, a tree diagram of the menu, an analogy which likened a search to moving through a department store, and a both a map and an analogy. Subjects performed eighteen searches on one day and then performed another eighteen searches twenty-four hours later. In analyzing the data, they found that on day two, subjects in the "analogy only" group were the least affected by increasing task difficulty. Webb and Kramer attribute this to the idea that spatial maps become less useful as menus become larger (i.e., more difficult).
Half of the aforementioned instructional research on menu selection have found significant differences between method of training and menu search performance, but at best these findings are controversial. Sometimes studying a tree diagram results in the best performance and in other cases studying an analogy results in the best performance.

Another point is that any differences that were found were only found after the subjects had completed more than half of their searches. In the case of Billingsley's research, significant differences occurred after the subjects had completed seventy-five percent of the required searches. These data would suggest that not only were the subjects getting used to the menu's structure, but perhaps subjects were also getting used to the type of instructional material they had studied. Perhaps if the subjects could be provided with tailored instructional material, the "break-in" time subjects need in order to acquire a working understanding of the information would be shortened. Strategies concerning how to tailor instructional material will be offered later.

Spatial Visualization Ability and Menu Search

Navigating a hierarchical menu is a spatially demanding task. One must mentally keep track of one's present position in terms of current level as well as relate that position to the starting point and the hypothesized location of the target. With these demands placed on the spatial abilities and the working memory of the user, it is easy to imagine users getting lost in large, unfamiliar menu systems. The following research has found two components of spatial ability that are related to performance in a menu system. The first component is spatial memory.

Spatial memory is defined as "the ability to remember the configuration, location, and orientation of figural material" (p. 109, Ekstrom, et al, 1976). Egan and Gomez (1985) found that spatial memory was a significant predictor of the subjects' performance when using a text line editor. Spatial memory predicted performance measures such as time to complete a search and number of errors made. This is a logical finding since one's memory of the location of an item in a long passage of text would affect how fast one can find that item and how error-free one's search would be.

The second component of spatial ability that was found to be a predictor of performance in a spatially demanding computer-based task was spatial visualization ability. Spatial visualization ability is defined as "the ability to manipulate or transform the image of spatial patterns into other arrangements" (p. 173, Ekstrom, et al, 1976). Research by Vicente, Hayes, and Williges (1987) examined how subjects with different spatial visualization abilities performed in a hierarchical file system. They found that subjects with low spatial visualization scores took longer to locate a target and exhibited behaviors that
led the researchers to conclude that the subjects were lost. Again, the idea that this component of spatial ability could relate to performance in a spatially demanding computer-based task, this time involving a hierarchical menu structure, is logical. Subjects who are able to manipulate the spatial pattern of their path should be better able to roam around inside a menu system while being less likely to get lost because they are able to retrace their steps.

The Present Study

Considering the above findings, it seems likely that providing subjects with instructional material that either complemented or supplemented their spatial visualization ability could have accelerated the learning process. More specifically, subjects with higher spatial visualization abilities might learn a menu structure more quickly if they were provided with spatially oriented learning materials such as a tree structure map of the menu structure. One might suggest that high spatial subjects would learn a menu system even faster if they were given both a map and an analogy--the more information about the menu, the better. It is hypothesized that high spatial subjects are able to integrate the information presented from the map so efficiently, that any added information the analogy might provide will not significantly affect performance.

As in the case of the high spatial subjects, subjects with lower spatial abilities might also have study aids which provide information optimally. Low spatial subjects do not perform well in hierarchical menu environments, as demonstrated by Vicente, Hayes, and Williges. This is probably due to the fact that they cannot build and store relations among items that are presented in a spatial fashion, as can high spatial subjects. Hence, low spatial subjects may need information presented to them in a less spatial manner--by way of an analogy. While low spatial subjects will need a map of the menu to learn specific information about the structure and rules of the menu system, they might also need the inner workings of the structure explained to them in a non-spatial manner. Hence, the need for a non-spatial analogy.

The fact that the low spatial visualization subjects gather their information from two different sources should put the low spatial subjects at a disadvantage to the high spatial subjects. The high spatial subjects should be able to integrate the rules/structure of the menu with items from the menu, as presented in one source: the map. On the other hand, the low spatial subjects will be forced to glean organizational rules of the menu from one source (a map) and then use another source (the analogy) to help in figuring out how the items and categories interrelate.
Specific Hypotheses

The above considerations lead to the following hypotheses that will be tested:

1) It is hypothesized that providing subjects with a certain level of spatial visualization ability with study materials that complement their level of ability will result in the subjects finding targets using fewer moves and in a shorter time than subjects who were not provided with such materials.

2) Providing subjects with complementary study materials should also result in an earlier peak in performance because the subjects need no time to acquire a working knowledge of the information contained in their study materials.

Independent Variables

The type of study material was varied across four conditions to provide all combinations of interest of map and analogy as well as providing a control condition where subjects studied information that did not concern the menu structure. The four study condition were:

(1) The Map study group was given a diagram that represented the structure of the hierarchical menu in a tree structure.

(2) The Analogy study group was given an analogy which compared searching the hierarchical menu to shopping in a mall. The analogy studied by subjects in Webb and Kramer (1988) likened searching a menu to shopping in a department store. Therefore, likening menu search to shopping in a mall is expected to be a useful analogy.

(3) The Combined study group was given both the diagram of the tree structure and the analogy which compared searching a hierarchical menu to shopping in a mall. Subjects in this study condition were allowed to divide their time between the map and the analogy in any fashion they wanted.

(4) The No Materials study group was given an article on Enrico Fermi to read during the allotted study period.

The second factor of interest was spatial visualization ability. Subjects were blocked on this factor into one of four groups: (a) High, (b) Medium High, (c) Medium Low, and (d) Low Spatial Visualization Ability. Assignment of subjects to different groups was performed according to the scores they received on a pre-test.

As stated previously, lower spatial visualization subjects are expected to perform significantly higher in the Combined study condition than any other condition because they will be provided with the specific information contained in the map while also having the map explained to them by way of an analogy. Higher spatial visualization subjects are
expected to perform equally as high in any condition which provides a map (the Map and Combined study conditions) because the information provided by the analogy will be superfluous--they will have already understood the relations between the items in the map simply by studying the map.

The third factor of interest was the effect of familiarity with the hierarchical menu. Four levels of familiarity were examined by breaking the forty searches into four blocks containing ten searches each. The subjects' expertise is assumed to have increased with an increase in block number. In other words, as the subjects completed more blocks of searches, their experience with the menu system increased. The repeated measures design which results from these three factors is graphically presented in Figure 1.

Insert Figure 1 about here

Dependent Variables

To examine the effects of these three factors, seven dependent measures were assessed. They were:

(1) Search Time—The time from the first path selection to the time of locating the target. This is a measure of overall performance. Search time includes time spent actively searching, time spent thinking, and time spent searching off target.

(2) Total Number of Moves—This number includes every choice a subject made in a particular search including any redirections and searches in the wrong part of the menu. This is also a measure of overall performance.

(3) Traversal Errors—The number of nodes a subject deviated from the correct path at the time of the first redirection of search. Given the menu structure employed, the traversal errors can range from a small error of magnitude, 1, to a large error of magnitude, 3. Traversal Errors are thought to provide a measure of effectiveness of a particular method of instruction.

(4) Total Starts—The total number of starts subjects issued in a block of ten searches. This was used to provide a rough measure of the degree to which subjects were feeling lost in the menu. This is considered a measure of how lost subjects felt because only by returning to start, can subjects be absolutely sure of where they are and it is expected that subjects who return to start many times are not quite sure of where they are when they are inside the menu system.
(5) Total Previouses— The total number of previouses a subject issued in a block of ten searches. This was considered to represent the degree of confidence a subject felt in navigating within the hierarchical menu structure.

(6) Subjective Rating Scales— The subjects rated "Ease of learning the menu structure" and "How lost did you feel" on unsegmented rating scales.

(7) Menu Representation— Subjects drew their mental representation of the hierarchical menu on a blank piece of paper. Items of interest were the number of correct menu items recalled and the type of representation. Type of representation concerned how subjects represented the menu structure; in a spatial or non-spatial fashion.

Method

Subjects

Sixty-four subjects were recruited from the University of Maryland subject pool which is comprised of students taking introductory psychology courses. Four subjects were randomly assigned to each cell in the Study Method by Spatial Visualization Ability matrix. In each cell, one subject was assigned to each of one of the four counterbalanced orders of blocks. Subjects were screened for prior computer use and those with more than three computer oriented classes were not accepted. Subjects received extra class credit for their participation in the study.

Materials

A 3x3x3x3 hierarchical menu was constructed using the HyperCard™ program for the Apple Macintosh. The menu contained information about three colleges within a hypothetical university and the various departments and classes that each college had to offer (see Figure 2). At the first level of the menu structure, subjects were presented with colleges within a hypothetical university. The three colleges were The College of Arts and Humanities, The College of Engineering, and The College of Behavioral and Social Sciences. After choosing a specific college, subjects progressed to the second level where they were presented with three "Departments" in the college. For example, in The College of Behavioral and Social Sciences, the three Departments were Psychology, Sociology, and Government and Politics. While in the second level, subjects were also given the option to return to the level of the colleges, or execute a "Back to the Top." Executing a "Back to the Top" was considered a "Start" in analyzing the subjects' performance. When subjects progressed to the third level after choosing a department, they were presented with three broad topics, or "Curriculum", that were offered within that particular department. In the third level, subjects were offered three choices of curriculum such as Clinical
Psychology, Counseling Psychology, and Social Psychology as well as the options to return to start or to go "Back One Step" to the level before their current level; in this case the second level. This stepping back was considered a "Previous" in the analysis of the subjects' performance. After choosing a curriculum, the subjects proceeded to the fourth and final level, the level of the "Class." This is where subjects searched for the target. At the level of the class, the subject would also have been able to execute a "Back to the Top" or a "Back One Step" command if none of the classes was the target class. Refer to Figure 3 for examples of how the data base was presented to the subject.

Insert Figure 2 about here

Insert Figure 3 about here

The Map study material was created by making a tree diagram of all the nodes in the course catalog and was presented to subjects on a single pieces of paper, as depicted in Figure 2. An analogy was created that likened searching a hierarchical data base to going shopping in malls. Refer to Appendix 1 for a transcript of the analogy.

The subject's spatial visualization ability was measured using the VZ-2 from Ekstrom, French and Harmon's 1976 Cognitive Test Battery. The VZ-2 is a paper folding test which depicts a piece of paper being folded four times and having a hole punched through it after the last fold is made. The subjects are asked to select, from four alternatives, the correct drawing of what the piece of paper looks like when it is unfolded.

Subjects were assigned to groups (high, medium high, etc.) on the basis of their score on the VZ-2. Scores required for the four groups were based on a sample of 505 VZ-2 scores of University of Maryland undergraduates. This sample had a mean of 13 and a standard deviation of 3.86. Assuming a normal distribution, the first quartile falls within scores of 1-9, the second within scores 10-12, the third within scores 13-15, and the fourth within scores 16-20. For purposes of the experiment, the first quartile was called "Low", the second was called "Medium Low", the third was called "Medium High", and the third was called "High."
Two rating scales were created to measure subjects' responses to the questions concerning how lost they felt and how hard the menu was to learn to use (See Appendix 2).

Four variations of the program were created wherein the four blocks of ten trials each were presented in different orders with Block A appearing once in all four positions (first, second, third, or fourth). Blocks B, C, and D were also counterbalanced across all four positions so as to eliminate any effects of block difficulty. Refer to Appendix 3 for a listing of the searches contained in each block.

Procedure

Subjects were first given six minutes to complete the VZ-2 cognitive test and it was scored in order to assign the subjects to the appropriate level of spatial visualization. No subjects were not given information about the nature of their scores. Once blocked, subjects were randomly assigned to a study condition. After being randomly assigned to a study condition, subjects were read general information about the experiment and their specific study condition (see Appendix 4). They were then given five minutes to study their instructional material and were told that at the end of the experiment, they would be asked to reproduce or answer questions about what they had studied. Then, all subjects performed five searches in a sample hierarchical menu structure in order to learn about the use of a mouse and how to press buttons in the Hypercard program. After subjects understood how to use the mouse and the "Back to the Top" and "Back One Step" functions, they began the experimental trials.

After completing all forty searches, subjects were asked to use the rating scales to respond to the two subjective questions concerning the ease of learning the structure of the database and the degree to which they felt lost while searching the database. They were also asked what Enrico Fermi's dropped at the nuclear test site (torn up pieces of paper). The question concerning Enrico Fermi was asked to determine whether the subjects in the None study condition had actually studied the article. After answering these questions, subjects were provided with a blank sheet of paper and something to write with and were asked to reproduce the menu structure and the information contained within it to the best of their ability. After producing their representation of the menu structure, subjects were debriefed, thanked, and given their experimental credit voucher.

Results

Four aspects of the subjects' performance were examined using a split plot design with repeated measures. The aspects of performance were: (a) difficulty of locating the
target, (b) redirection strategies the subjects employed in their searches, (c) errors as measured by deviations from the path which lead to the target, and (c) knowledge about the structure and contents of the menu. Subjective measures concerning ease of learning to use the menu system and how lost the subjects felt while inside the menu were also administered.

Locating the Target— Number of Moves and Speed

The first aspect of performance concerned how difficult it was for subjects to locate the targets. This was measured by two dependent variables: search time and number of moves. The number of moves per block was calculated by adding up all the menu selections subjects made while searching for targets in a block of ten searches. Search time was calculated by adding up the total time it took subjects to locate ten targets in a block of ten searches.

The number of moves subjects needed to locate targets was affected by all three factors of interest. The type of material the subjects studied had a significant effect on the number of moves (F (3, 48) = 5.23, p < .05) as shown in the top panel of Figure 4. Tukey's HSD was used to determine which methods of study were significantly different at the .05 level. Subjects who studied the map or the map and analogy made significantly fewer moves than subjects who read the article about Enrico Fermi.

The level of spatial ability subjects possessed also affected the number of moves subjects made (F (3, 48) = 3.66, p < .05). The high spatial subjects made significantly fewer moves than the low spatial subjects as indicated by Tukey's HSD and as shown in the middle panel of Figure 4.

The amount of familiarity subjects had with the menu system affected how many moves they needed to locate the targets with the first block of searches requiring significantly more moves than the third and fourth blocks (F (3, 144) = 7.29, p < .05), as illustrated in the bottom panel of Figure 4. There were no significant interactions between any of the factors of interest.

Insert Figure 4 about here

The second measure of search difficulty, search time, was affected by spatial ability with the low spatial subjects taking a longer time to find targets than high spatial subjects (F (3, 48) = 3.639, p < .05) as shown in the top panel of Figure 5.
The time subjects took to find the target was affected by a block by study interaction \((E(9, 144) = 2.047, p < .05)\) as illustrated by the bottom panel of Figure 5. No other interaction was significant.

Search time was also positively correlated with number of moves \((r(62) = .85, p < .0125)\). (Significance of the four correlations was tested at the .0125 level due to a Bonferroni adjustment.)

**Redirecting Strategies**

The second aspect of performance concerned the subjects' redirecting strategies once they were within the menu and decided to redirect search. Redirecting strategies were measured by two dependent variables: "previouslys" and "starts." Again, the values of these dependent measures were obtained by adding up the number of starts and previouslys that subjects made in blocks of ten searches. The number of previouslys subjects made was influenced by the type of material they studied with subjects who studied a map of the menu structure making fewer previouslys than subjects who studied the analogy or the article about Fermi as illustrated by the top panel of Figure 6 \((E(3, 48) = 3.72, p < .05)\). The number of previousy makes was significantly correlated with number of moves \((r(62) = .87, p < .0125)\).

The number of starts was affected by the amount of experience subjects had with the menu \((E(3, 48) = 10.92, p < .05)\). Subjects made more returns to start in the first block than in the second, third, and fourth block of searches as shown in the bottom panel of figure 6. The number of starts was not significantly correlated with number of moves \((r(62) = .29, p > .0125)\).

**Deviations from the Optimal Search Path**

The third facet of performance measured the magnitude of subjects' deviations from the correct path at the time they decided that they were looking in the wrong place. This variable is called "traversal errors." The number of traversal errors was calculated by counting the number of nodes a subject would have to traverse in order to get back on the right track and search for the current target. Traversal errors were calculated at the time of
the first redirection. The magnitude of the subjects' traversal errors was affected by the type of material they studied with subjects who studied a map of the menu making smaller traversal errors than subjects who studied either the analogy or the article about Fermi as shown in Figure 7 ($\bar{E}(3, 48) = 6.49, p < .05$).

The number of traversal errors subjects made was significantly correlated with the number of moves subjects made ($r(62) = .76, p < .0125$).

**Knowledge about the Menu Structure**

The fourth aspect of performance was concerned with what the subject had learned about the structure of the menu system and the information contained within it. Two variables were used to examine the subjects' knowledge of the menu. The first variable measured the total number of items subjects were able to recall and write down in their rendition of the menu structure. This is "Items Recalled." The second variable examined the subjects' rendition of the menu structure. Was their rendition spatial in nature with features of a tree diagram and therefore a more accurate representation of the menu structure, or was it merely a listing of nodes and labels?

The number of items recalled was affected by what materials the subjects were given to study ($\bar{E}(3, 48) = 3.2, p < .05$). As shown in Figure 8, subjects who studied both an analogy and a map of the menu were able to recall more labels from the menu structure than subjects who studied the article about Fermi.

Differences in the style of subjects' renditions of the menu were examined by using a Chi-Square contingency test. First, the effect of study materials was examined. Differences in rendition style were apparent. Subjects who studied a map or both a map and an analogy tended to draw a map of the menu rather than merely list the labels. Subjects who were given the analogy or the Fermi article tended to make lists or outlines of labels rather than organize them in a hierarchical fashion. This effect was significant at the .05 level. Refer to the top panel of Figure 9 for a graph depicting the differences in
rendition style. The Chi-Square contingency test is presented in the bottom panel of Figure 9.

The effect of spatial visualization ability on the subjects' rendition of the menu was also examined using a Chi-Square contingency test. The data were split into two groups around the median. The group falling below the median was the "lower" spatial visualization group and the group lying above the median was called the "higher" spatial visualization group. The Chi-Square test was not significant at the .05 level. Hence it appears that increases or decreases subjects' spatial visualization ability has no effect on what kind of mental model of the menu they made. Refer to the top panel of Figure 10 for a graph of the data. The Chi-Square contingency test is presented in the bottom panel of Figure 10.

In the above Chi-Square analyses, the four study conditions were combined into two groups. The first group contained the map and both study conditions. The second group contained the analogy and Fermi study conditions. The study conditions were combined in this fashion for a variety of reasons. The first reason is that the pattern of results for subjects in the map and both study conditions is very similar as is the pattern of results for subjects in the analogy and Fermi groups. The second reason is that subjects in the both study condition, tended to spend the majority of their time studying the map, thus making the map and both study conditions comparable. The analogy and Fermi study conditions were combined because it was thought that the hierarchical organization implied by the analogy was also implied by the task domain. That is, college students know universities are organized hierarchically and they can transfer that knowledge of hierarchies to the menu system. Hence, subjects who studied the analogy took a "shopping mall hierarchy" into the task with them and Fermi subjects took a "university hierarchy" into the task domain. Since both groups of subjects took some sort of hierarchical orientation with them into the task, they were considered to be comparable.
Subjective Measures

Two rating scales were presented to the subjects when they finished their searches. The first concerned a difficulty rating for learning to use the menu program. The second concerned a self-report of how lost the subjects felt while they were searching within the menu structure. Neither measures were affected by what the subjects studied about the menu structure or what level of spatial ability they had.

Discussion

Main Hypotheses

The hypotheses proposed concerned the effects of the interaction between the materials a subject studied and the level of spatial visualization ability a subject possessed. As previously discussed, it was hypothesized that given a certain level of spatial visualization ability, subjects would learn most efficiently if provided with study materials which complement their spatial visualization ability. Specifically, high spatial subjects would learn most efficiently from a map because they would be able to extract more information from a map than other study material combinations. Conversely, low spatial subjects would not be able to extract enough information from just a map; they would need the map explained to them by way of an analogy. From this line of reasoning, the following hypotheses were proposed about the study materials by spatial visualization interaction:

1) Providing subjects with efficient, complementary study materials would result in lower search times and fewer moves to find the target.

2) Subjects who studied materials which complemented their spatial visualization ability would reach their peak performance in earlier trials than subjects who did not study materials which complemented their spatial ability.

Neither of the two main hypotheses were supported by the study by spatial interaction for total moves and search times. Three reasons are suggested for why the interaction of interest failed to reach significance.

The first reason suggested concerns a problem with the sampling of subjects and their motivation. About 75% of the subjects were run during the last three weeks of school during the spring semester of 1989. The other 25% were run during the first two weeks of the fall semester of 1989. On four occasions in the Spring, subjects got exasperated during the experiment and made verbal ejaculations along the lines of "This is stupid" and "When am I going to be done?" None of the fall subjects made negative comments during the experiment. This leads one to believe that the subjects who ran toward the end of the
spring semester could very well have fit the stereotype of the below average, low motivation student who is volunteering to be a subject simply to get extra credit so as not to fail PSYC 100.

Another point supporting the idea that the spring and fall subjects were from two different populations is that of the twelve subjects collected in the fall, eight of them were high spatial subjects. The other four subjects were divided between the medium high and low visualization groups. This implies that the majority of the lower spatial visualization subjects were run at the end of the spring semester when motivation is assumed to be at a premium among subjects. Considering the idea that the materials that were assumed to aid low spatial subjects the most were also the most tedious to study (the analogy or the map and analogy), the low visualization subjects were doubly penalized. They were double penalized because not only were the low visualization subjects lacking the innate abilities to perform well in the menu, but they probably didn't study the very materials that were designed to help them. This probably lead to poor overall performance regardless of the material studied.

The second reason being offered for why the spatial visualization by study material interaction was not significant concerns a design flaw about spatial visualization. Spatial visualization was described as being "the ability to manipulate or transform the image of spatial patterns into other arrangements" (p. 173, Ekstrom, et al, 1976). Spatial memory was described as being "the ability to remember the configuration, location, and orientation of figural material" (p. 109, Ekstrom, et al, 1976). Spatial visualization was chosen as the ability of interest because it was assumed that the ability to manipulate spatial patterns was important for high performance in a hierarchical menu. This decision was made based on the findings of Norman and Butler (1989) and Vicente, Hayes, and Williges (1987). Norman and Butler found that a significant negative correlation existed between scores on the VZ-2 and the number of moves and the number of deviations from the best search path while Vicente, et al found subjects with low spatial visualization scores took longer to locate targets. Hence, spatial visualization ability was chosen because it was so firmly associated with performance in hierarchical menus.

Spatial memory was not a factor of interest because the design was becoming very large and because the present study was concerned with how well subjects could manipulate information they could remember. Since spatial memory is concerned only with how well subjects are able to remember figural material, it was thought to not be important to the subjects' ability to manipulate the information presented in the various study materials. Here is where the design flaw might have appeared. If the level of spatial
memory governs how well subjects remember spatial material, one might argue that even high spatial visualizers could not manipulate the material if they were low in spatial memory ability and couldn't remember what they had studied. This line of reasoning is supported by the fact that subjects' level of spatial visualization ability did not affect how many items they were able to recall from the menu.

The present study was concerned with presenting learners with study materials that would complement some ability of theirs. Perhaps it is possible to provide learners with study materials that will enable them to assimilate information optimally, but it seems that such materials should be distributed on the basis of the learners' ability to remember them. Perhaps the study materials should be distributed on the basis of spatial memory. Such a suggestion paves the way to a logical extension which might include a model of performance in hierarchical menus which is comprised of spatial memory ability and spatial visualization ability.

The third reason suggested for why the interaction failed to reach significance concerns the analogy which may have been implied by the task domain. It seems likely that once subjects in any study condition saw the choices at the top node, "The College of Engineering", "The College of Arts and Humanities", "The College of Behavioral and Social Sciences", a hierarchical organization, one of a university, would have been implied. The hierarchical organization of a university was probably a very familiar concept to the subjects since they were all college students and as such was probably quite apparent. Not only was it probably a familiar concept to the students, but it also quite accurately described the organization of the menu—the menu was organized like a university. Once the non-analogy subjects made the inference that the menu was indeed hierarchical in nature, they knew as much, or more if they were in the map condition, than did the subjects who studied the analogy.

From the above argument, it seems likely that all subjects, regardless of study condition, were exposed to an analogy which implied some sort of hierarchical organization. If all subjects had an analogy made available to them, then the effects of study material may have been confounded, thus eliminating any interaction between study material and spatial visualization that may have occurred.

**Study Material Main Effects**

Although the interaction hypotheses were not supported, the results for the main effects are logical and consistent with previous research. Subjects who studied either a map or both a map and analogy made fewer moves in searching for targets than did subjects who studied either the analogy or the article about Fermi. Hence, the map seems
to be the crucial study material. This finding makes sense because the map provides the subject with information about exactly what is in the menu and where it is located. An analogy, on the other hand, only tells the learner what the menu is like and gives no specific information about rules of organization or items within the menu.

Subjects who studied the map made the fewest deviations from the best search path as compared to subjects who studied the analogy or the Fermi article. Again, the superiority of the map as the best study material is supported. Deviations from best search path were measured by how far off the correct path the subject went before deciding to redirect search. The map seems to have given subjects the best overall impression of the rules by which items were organized in the menu system. Subjects didn’t stray as far from the correct topical area before deciding to redirect search.

The magnitude of deviation from the search path was significantly correlated with total number of moves, but this correlation does not imply a causal relationship. In this case, it seems likely that just as a more complete representation of the menu structure will result in fewer moves, so it will result in a more timely decision to redirect search. This idea of map study subjects having a more complete representation of the menu is supported by the significant results for method of study and number of items recalled.

Subjects who studied the map made fewer previouses than did subjects who studied either an analogy or the Fermi article. This difference is probably due to the fact map subjects made fewer moves in general rather than reflecting any particular abilities or lack thereof. This argument is supported by the significant correlation between number of moves and number of previouses.

The number of items subjects included in their representations of the menu was affected by type of material studied; Subjects who studied the map were able to recall more items than were subjects who studied the Fermi article. Subjects who studied the map or the map and analogy also rendered spatial representations of the menu structure while subjects who didn’t study maps rendered list-like representations. Again it seems that maps, more than any other study material or combination thereof, lead to more accurate representations of menu structures regardless of the subjects’ spatial abilities.

Spatial Visualization Main Effects

Effects of spatial visualization ability were generally the result of gross differences between low spatial subjects and high spatial subjects. In other words, subjects had to be severely lacking in spatial visualization ability for that deficit to affect their performance. Subjects with high scores on the VZ-2 made fewer moves in locating the targets and took less time to locate targets than did subjects with low scores as illustrated by Figure 11.
These differences in number of moves and search times are consistent with the findings of Norman and Butler and Vicente, Hayes, and Williges.

--------------------
Insert Figure 11 about here
--------------------

**Repeated Measure— Practice Effects**

As might be expected, subjects needed fewer moves to locate targets after they had gained experience with the menu system. The effect of experience was evidenced in the third and fourth blocks.

Another effect of experience appeared in a block by study material interaction, as illustrated in the bottom panel of Figure 5. In the first block of searches, subjects who studied the Fermi article took about one and a half times as long to find targets as subjects in the other study conditions. But in blocks two to four, the Fermi subjects performed more and more like the analogy subjects, while the subjects who studied the map or both the map and analogy performed almost identically. By trial four, the initial differences between Fermi subjects and the others had diminished drastically although the Fermi subjects still had the longest search times. It appeared as if the long search times for Fermi subjects in block one were as a result of their having to formulate the rules of the menu system during their first block of searches. The other subjects had already gotten most of the rules of the menu from their study materials, but the Fermi subjects had to figure out the rules of the data base as they went along; a disadvantage they never quite overcame, as evidenced by their high search times.
Conclusions

It appears that one cannot affect performance in a hierarchical menu by tailoring study material to the learner's spatial visualization ability—a map is the best study material for all subjects. Furthermore, it does not appear that differences in menu search performance exist between fine differences in spatial visualization ability. Hence, in selecting people to work with hierarchical products, one need not worry about fine differences in spatial visualization ability among potential employees. As long as people are not severely lacking in this ability, they will perform well.

Applications of Results and Suggestions for Future Research

These results should be good news for employers, software users and software designers. Employers do not need to be concerned with fine differences in their employees' spatial visualization ability so as to properly assign the best suited workers to hierarchical tasks. Instead, they should be concerned with gross deficits in spatial visualization ability and what kind of instructional materials they provide to their workers. Software users do not have to read lengthy descriptions and analogies that compare hierarchical menus to shopping malls and department stores. Instead they should be provided with maps in order to see exactly what is contained in the menu and how it is arranged. Better still, software designers do not need to write the lengthy descriptions and analogies and worry about tailoring instructions to the characteristics of the user. Rather, they can put all their effort into one approach, namely, a good map of the menu structure.
References


Figure 1: A graphic representation of the experimental design.
Figure 2: The structure of the hierarchical menu and its contents. Also the study prop used in the Tree study condition.
Figure 3: The screens at all four levels of the hierarchy from topmost node (at the top) to the bottommost node (at the bottom).
Figure 4: The mean number of moves per subject combined across groups other than type of material studied (top panel), spatial ability (middle panel), and block (bottom panel).
Figure 5: Mean search times per subject combined across groups other than spatial ability (top panel). Mean search times per subject presented as a function of study condition across blocks (bottom panel).
Figure 6: The mean number of previouses per subject combined across all groups except study material (top panel). The mean number of starts per subject combined across groups other than block number (bottom panel).
Figure 7: Mean number of deviations from the optimal search path, per subject, combined across groups other than study material.
Figure 8: Mean number of items recalled per subject from subjects' rendition of the menu structure combined across groups other than study condition.
Figure 9: Number of spatial/list renditions of the menu structure combined across groups other than study material (top panel). Data presented in Chi-square form (bottom panel).
Figure 10: Number of spatial/list renditions of the menu structure combined across groups other than spatial ability (top panel). Data presented in Chi-square form (bottom panel).
Figure 11: Scatterplots illustrating the trend of high spatial visualization subjects requiring fewer moves to find targets (top panel) and less time to locate targets (bottom panel).
Appendix 1

Later in this experiment you will be using a computer menu system to search for classes that are offered at a hypothetical university. A computer menu system is a convenient way of organizing information into groups so that it can be easily retrieved. The information you will be searching has a logical, unchanging structure and these instructions are being provided so that you will be as efficient as possible when you search through the menu system.

Searching a computer menu system is a lot like going shopping in an area with a lot of malls. When you're looking for classes, the computer will first ask you which of three colleges you think the class is in. This is like choosing which one of three malls you want to shop in. Each of the three malls has different merchandise to offer and if you are looking for a specific item, you'll choose the mall you think is most likely to have it. Just like different malls offer different merchandise, different colleges offer different classes.

After choosing the college you think offers the course you are looking for, you have to choose a school within that college. This is like choosing a store within the mall. Each store offers different merchandise just like each school offers different courses. After picking a school to search, you have to pick an area of specialization. This is like choosing a department in a store. Again, different departments offer different types of merchandise and different areas of specialization offer different classes. After choosing an area of specialization, you are shown the classes that are offered; this is like being shown the merchandise. So, in order to get to the merchandise, first you have to pick a mall (college), then you have to pick a store (school), and finally you have to pick a department in the store (area of specialization). Once you're in a department, you will see if it has what you are looking for (the class).

If the choices that a department offers you are not what you are looking for, you can do a few things. You can go "Back one Step" to look in a different department in the same store. You can go back two steps to look in different stores, but stay in the same mall. Or, you can go "Back to the Top" and change malls if you think that nothing else in the the mall you are in will fit what you are looking for.

Of course, any time you go "Back One Step", it's like walking out a door. So if you walk out a door when you're in a department, you are in a store and can choose from any department in the store. When you walk out a door of a store, you're in the mall and can choose from any store in the mall. When you walk out the door of a mall, you're in the parking lot and can go to any other mall. Whenever you go "Back to the Top" you immediately go to the parking lot.
Appendix 2

The Likert Scales for subjective rating of ease of learning
and feeling of being lost

Instructions: Please put a mark somewhere along the two lines below
that would indicate what your responses are to the two
questions. Please try to be as accurate as possible

Example: The question is "How much do you like green beans?"
Let's say you really like green beans. You'd respond
something like:

I don't like them at all

[Mark]

I like them a lot

I never felt lost

[Mark]

I frequently felt lost

2. How easy was it to learn how to use the menu system?

It was very hard to learn how to use

[Mark]

It was very easy to learn how to use
Appendix 3

A listing of the four blocks and the target classes they contained.

**Block A**
- Dostoevski— The Brothers Karamazov
- Chekov– The Three Sisters
- Kierkgaard's Philosophy
- Intelligence Testing
- Antennas
- Dam Management
- Intro to Public Administ'n and Policy
- Population Dynamics
- The Works of Salvador Dali
- Machine Design

**Block B**
- Shakespeare— The Tragedies
- Tchaikovsky— The Symphonies
- Construction Equipment and Methods
- Personal Growth Group
- Alcohol Abuse and the Parent
- Frédéric Chopin— The Waltzes
- Computer Structures
- Automatic Controls for Factories
- Rodin and His Sculptures
- American Foreign Relations

**Block C**
- Highway and Airfield Paving Methods
- Historical Emergence of the Afro-American
- Thermodynamics
- Michelangelo— The Renaissance Man
- Antonin Dvorak— The Man and His Music
- The Judicial Process
- History of the Netherlands
- Hegel's Philosophy
- Electromagnetic Theory
- Group Dynamics

**Block D**
- Dream Analysis
- Divorce in the 1980's
- Taking the Census
- Churchill— The War Years
- Interpersonal Relations
- The Religious Writings of Martin Luther
- Dynamics II
- El Greco and His Art
- The Works of Cervantes
- Russian Political Thought
Appendix 4

Instructions to the Subjects for the Four Study Conditions

General Instructions: In this study you will be asked to do two tasks. The first task is completing a test which measures spatial visualization ability, the VZ-2, and the second task is performing a number of searches for class names in a computer database system. Spatial visualization ability is the name we give to one's ability to "see" shapes in one's head. Before you perform the database searches, you will be asked to complete a short paper and pencil test which will give us a rough measure of your spatial ability. Do you have any questions about what is required of you in this experiment? (Ask subject to sign Informed Consent Form.)

Please read this page of instructions for the VZ-2. They should seem very simple and straightforward. If you have any questions, let me know. When you understand the instruction, let me know and you may begin the first half of the test. You will have three minutes to complete each half of the test. Don't worry if you don't finish; we deliberately cut the time short so people would have a hard time finishing.

Instructions to Subjects in the Tree Study Condition: The main part of the experiment will consist of your looking up class names in a database. Here is a representation of the structure of the database (show tree structure to subject). Please study this diagram for the next five minutes. At the end of the experiment you will be tested on what you can remember about this diagram.

Instructions to Subjects in the Analogy Study Condition: Here is an analogy of how the database you will be using is constructed. This analogy also offers insight into relations between the information which is stored within the database. Please read the following passage for the next five minutes. At the end of the experiment you will be tested on what you can remember from this passage.

Instructions to Subjects in the Both Study Condition: The main part of the experiment will consist of your looking up class names in a database. Here is a representation of the structure of the database (show tree diagram to subject).

Now, please read the following passage and feel free to refer to the pictorial representation of the database to aid yourself in understanding the structure and interrelations of the database. You will have five minutes to study these materials. At the end of the experiment you will be tested on the amount of material you remember from the diagram and the passage.

Instructions to Subjects in the No Material Study Condition: Please read this article about Enrico Fermi for the next five minutes. At the end of the experiment you will be tested on the amount of material you remember from the passage.

General Instructions: In using this computer and searching for classes, there are three things you need to know about. The first is the mouse which allows you to point to objects on the screen by directing this arrow. When the arrow is on top of what you want to select, simply click this button and your selection will be made. When you are inside the database you will see three types of "buttons" that you could select. They all move you through the database, but some move you towards the class you are looking for and some move you back to the start. Buttons like this (gesture to Sample Program) move you toward the desired class. These buttons marked "Previous" and "Return to Start" return where you were before you made your last choice and return you to the level of the colleges, respectively. These two buttons are used to redirect your search if you feel that you may be looking in the wrong place. Are there any questions about how to use these buttons?

You will complete forty searches. When you are performing your searches, try to search as quickly and efficiently as possible. Good luck.