Window Control Strategies
for Hypertext Traversal:
An Empirical Study

Jacob (Kobi) Lifshitz† and Ben Shneiderman
Human-Computer Interaction Laboratory
Department of Computer Science
University of Maryland
College Park, Maryland 20742

Abstract

Nowadays, larger and higher resolution screens supporting multiple windows are widely available. They are often used to traverse hypertext databases for fact retrieval, education or casual browsing. This paper describes a study comparing two methods of window control strategies for article placement in four, equal sized and tiled windows (By tiled we mean non-overlapping, non-movable and non-resizable windows, which together cover the entire screen.) The first placement strategy - Automatic Panning (AP) - was controlled by the system software and the second placement strategy - User Controlled Placement (UCP) - transferred control to the users. The task required subjects to answer questions about the content of a hypertext. Statistical analysis shows that subjective preference was significantly higher and amount of backtracking was significantly lower for the user controlled placement. Speed differences did not prove to be significant, although there was consistent tendency favoring the automatic panning strategy.


† Current address: Georgetown University Hospital, Hospital Information Systems, 3800 Reservoir Road N.W.,
Washington, D.C. 20007.
Window Control Strategies for Hypertext Traversal: An Empirical Study

Jacob (Kobi) Lifshitz† and Ben Shneiderman*
Department of Computer Science &
Human-Computer Interaction Laboratory
University of Maryland

Abstract
Nowadays, larger and higher resolution screens supporting multiple windows are widely available. They are often used to traverse hypertext databases for fact retrieval, education or casual browsing. This paper describes a study comparing two window control strategies for article placement in four, equal sized and tiled windows (by tiled we mean non-overlapping, non-movable and non-resizable windows, which together cover the entire screen.) The first placement strategy - Automatic Panning (AP) - was controlled by the system software and the second placement strategy - User Controlled Placement (UCP) - transferred control to the users. The task required subjects to answer questions about the content of a hypertext. For each question, the system recorded the length of the search path, usage counts of the index and backtracking, and the time it took to solve the problem. Statistical analysis shows that subjective preference was significantly higher and amount of backtracking was significantly lower for the user controlled placement. Speed differences did not prove to be significant, although there was a consistent tendency favoring the automatic panning strategy.

1. Introduction
In the past few years, numerous hypertext systems have been developed for a variety of computer platforms utilizing the most advanced technologies. They include very high resolution display devices which facilitate the use of multiple window systems, embedded menus, and very high capacity storage facilities such as CD-ROM. Hypertext has emerged as a powerful tool and has been utilized in a variety of applications, including on-line manuals, on-line help instructions, company policy guides, interactive encyclopedias, specialized reference systems, and shared document universes.

Although hypertext is perceived as an elegant approach for linking and integrating information with relative ease of use, nonetheless it introduces new challenges for users with regard to relevance, orientation, navigation and the navigation strategies. Relevance refers to the appropriateness of the material being retrieved in relation to a task domain or interest. Orientation refers to the user’s awareness of the current system state in relation to other possible states. Navigation refers to the ability to move from one system state to another. To maximize interaction results, it is important to navigate efficiently through the nodes and links of the hypertext. The longer the expected path, the more important it is to develop a navigational strategy which takes into account relevance and orientation, and which can be modified flexibly as users’ current needs change.

† Current address: Georgetown University Hospital, Hospital Information Systems, 3800 Reservoir Road N.W., Washington, D.C. 20007.
* Ben Shneiderman is a member of the University of Maryland Institute for Advanced Computer Studies

2. Window Systems and Hypertext

The match between window systems' interface features and data in hypertext form is apparent. As Nelson's idea of hypertext suggests [Nels 87a], hypertext is made up of "discrete, chunk style pieces" (p. 134). Each of those chunks can be naturally placed within its own window, representing the hypertext as a collection of windows, each having a well defined content. This approach, with minor variations, was adapted in many implementations such as NoteCards, Intermedia, Guide and others.

Despite the intuitive acceptance of the benefits of using multiple windows for hypertext environments, a legitimate question can be raised about the validity of these assumptions, challenging the merits of a window system when compared with conventional single window displays. It was suggested by Bury et al. [Bury 85] that the benefits of multiple windows reduce the user's short-term memory load, thus freeing more cognitive resources to the user's task. In their study, Bury et al. compared a multiple window system to a single window system. Their analysis showed that overall performance was faster for the single window system. However, an analysis of the task-completion when broken down into component actions showed that time spent in the primary task domain was shorter when using the multiple window system. The remainder of the time was largely spent on window housekeeping.

Bly and Rosenberg [Bly 86] compared tiled and overlapped windows for matching graphical objects with corresponding text. In one treatment the objects and text elements were ordered while in a second treatment they were randomly distributed. Their conclusion was that "for a given level of expertise, tiled windows are better in both situations" (p. 104).

Both results suggest that users of overlapped window systems attend to two competing domains: the primary task domain and housekeeping overhead.

3. Browsing Strategies and Control

The current research stemmed from the effort of porting Hyperties from the PC environment, where hypertext nodes were displayed in a non-window environment, to high resolution, multi-window workstations. As our research on the SUN workstation progressed facilitating the use of a number of simultaneously open windows, the issue of article placement control (user vs. system) surfaced repeatedly.

Lifshitz [Lifs 89] suggested numerous possibilities for article placement strategies. Considerations were given for automatic placements vs. user controlled placements. Some automatic strategies for a fixed number of non-manipulable tiled windows included the panning of text by pages or articles, where the newly retrieved article was always displayed in a specially designated window. The rotation strategy is similar to panning except that new articles are placed over the least recently retrieved article. Some variations allowed new article selections to be made only from the most recent article while others did not limit the selection to any particular window. User control strategies included overlapped windows with a 1-1 correspondence between articles and windows. Another user controlled strategy suggested the use of fixed non-movable tiled windows; however, the choice of article placement was left for the user.

When we are dealing with control trade-offs, some people may argue that tighter system control may facilitate faster, more efficient and consistent interaction. It minimizes learning because of the limited interaction language that is necessary, and makes the application look more uniform. Others may argue that, in general, humans prefer to be in control, and that their performance deteriorates when they feel dominated. Some benefits of user control may include increased flexibility, predictability and satisfaction.

4. Reading comprehension task analysis

The component breakdown analysis of reading comprehension tasks for the two control
strategies that were chosen for the experiment is as follows:

Automatic Panning
1. Task conceptualization
2. Search strategy development
3. Article selection
4. Reading
5. Information evaluation

User Controlled Placement
1. Task conceptualization
2. Search strategy development
3. Article selection
4. Organizational strategy
5. Destination selection
6. Reading
7. Information evaluation

The extra steps that are part of the User Controlled Placement indicate that the advantage of added flexibility is associated with other trade-offs, as described earlier. While reducing some of the load in the task domain, new human resources have to be allocated for utilizing the User Controlled Placement. In the wake of the trade-off analysis between User Controlled Placement and Automatic Panning, it is interesting to find the balance point where the advantage of using one strategy gets shifted to the other strategy. In order to do so, it might be useful to look at the additional dimension of task classification.

An added dimension to control strategies is the task difficulty level. As the task becomes more open-ended, the necessary path length increases and the search becomes less obvious, and requires users to re-evaluate the relevance for each piece of information at the moment of retrieval. A User Controlled Placement, which allows users to keep relevant information independently of the path traversed, is believed to be better than any strategy that does not take into account the relevance of articles to the task at hand.

It is this hypothesis that led to the following experiment which examined the relationships of browsing performance with different window control strategies and different task difficulty levels.

5. Experiments

In order to test the hypothesis that increased user control can benefit hypertext users in complex retrieval tasks, it was necessary to narrow the scope of the study in order to eliminate as many interfering variants as possible. Therefore, the following assumptions were made:

1. Considering the results by Bly and Rosenberg, and Bury et al., and our desire to focus subjects' attention on the hypertext context, tiled windows were used which cover the screen entirely and which do not permit any movement or resizing operations. Thus, only two mouse operations were required: selection and confirmation.

2. Four windows were chosen because it was a large enough number to allow sufficient user flexibility while at the same time leaving the window size satisfactory.

3. Since the primary interest was to study the manipulability of the articles rather than the particular software features of Hyperties, it was decided to replace the definition with a pop-up window which prompted users for the destination window.

4. To eliminate bias due to different number of mousing steps needed for the two strategies, it was decided to display the pop-up window in all cases and require users to make a redundant confirming mouse click for the system control strategies. Although some may argue that this decision unnecessarily increases the bias in favor of the User Controlled Placement, it reduced the variability between the three experimental components.

5. In an effort to make the experiment more manageable, it was necessary to limit the number of strategies that were tested. Since panning-by-articles with a partially active background seemed to be the most powerful, it was chosen to represent the System Control strategies.

6. At the start of every task the system reinitialized the windows, so the position within the database due to earlier tasks would not introduce additional variance. This also guaranteed that
The Holocaust and Its Origins:

The Case of Austria

by

Marsha Rozenblit

Darlal Treat Courtney

ORGANIZATIONS: CHRISTIAN SOCIAL PARTY PAGE 2 OF 3

The Christian Social Party channeled this social protest into a full-fledged attack on Liberal political hegemony in Vienna. Promising those groups which hated the Jews to put an end to so-called Jewish "power," and downplaying such antisemitism among those who found it distasteful, Lueger dazzled the Viennese bourgeoisie with his oratory and led the Christian Social Party to victory in 1895. The Party maintained its political dominance in Vienna until 1920 when it was defeated by the Social Democrats. The Christian Socials did not enact any of their antisemitic promises during this period.

After its victory in Vienna, the Christian Social Party became the major bourgeois political force in Austria, and dominated politics in the Interwar First Republic. Christian Social leaders like Ignaz Seipel, Engelbert Dollfuß, and Kurt von Schuschnigg served as prime ministers of Austria.

PEOPLE: IGNAZ SEIPEL PAGE 1 OF 2

Born in Vienna in 1876, Ignaz Seipel was a Roman Catholic priest whose public career included five terms as chancellor (1922-24, 1926-29) of the Austrian First Republic. As the leader of the conservative, clerical Christian Social Party, Seipel dominated

Austro-fascism of the Heimwehr as convenient tools in the struggle against socialism. In the tradition of the Christian Socials, Seipel did not possess a pan-German political outlook, and he therefore worked hard to maintain and strengthen Austrian independence.

NEXT PAGE BACK PAGE

UNDO Next Question Question #4: Three Christian Socialists became leaders of Austria. When did they first enter office?

Answers: A. 1922, 1932, 1934

B. 1926, 1932, 1934

C. 1922, 1929, 1932

D. 1922, 1926, 1929

NEXT PAGE INDEX
the optimal distance for each task completion remained equivalent for all subjects.

5.1. Experimental Method
The study was implemented using an extension of Hyperties and was adapted to implement three different strategies. Two strategies, Automatic Panning and User Controlled Placement, were used to study the primary hypothesis while the third, a Single Window strategy, served as a statistical covariate to factor out individual differences between the subjects. Both the Automatic Panning and the User Controlled Placement implementations divided the screen into four tiled windows of equal size.

In all versions, each window was devoted to the display of a single article or the index and could display up to 20 lines per window. Selectable words in the text were highlighted by a bold font. Upon selection of such a key word, a pop-up window was displayed next to the selection which included an instruction for the appropriate version and a map representing the window layout on the screen. To complete the selection, subjects were required to select the appropriate window on the map. A small section at the bottom of the screen was used for managing the experiment and was not part of any window involved in the experiment.

5.1.1. Single Window
The Single Window implementation displayed all the articles in the same window at the top left of the screen. The rest of the screen remained blank and no other windows were displayed. The window was placed at the upper left corner of the screen to conform to the common reading strategy in English, which is from left to right and top to bottom.

5.1.2. Automatic Panning
The Automatic Panning implementation designated the lower right window as the active window. Upon completion of a selection process, newly selected articles were placed in the active window. Each subsequent selection pushed older articles in a zig-zag motion from the lower right window to the top left window, with the pop-up indicating the direction of the movement.

To indicate that only one window was active, only the lower right window on the pop-up map was clear and selectable. The other three inactive windows were grayed out, and were not selectable [Figure 1].

5.1.3. User Controlled Placement
The User Controlled Placement implementation allowed users to select any of the windows as the destination of the selected article. No order existed among the windows and the decision of where to place new articles was left entirely to the user. Subsequently, all the windows on the pop-up map were clear and selectable [Figure 2].

5.2. Database
Two databases were used in the experiment. A small recipe database was extracted from "The Silver Palate Cookbook" by Julee Rosso and Sheila Lukins and was used for the practice sessions. The recipe database included a total of 22 loosely linked articles. A larger database about the Holocaust in Austria, containing 106 tightly linked articles, was used for the experiment. The Holocaust database was written by Dr. Marsha Rosenblitt and Darla Treat Courtney.

5.3. Subjects
Twenty-three subjects volunteered to participate in the experiment. Approximately two-thirds of the participants were professional and knowledgeable computer users. Excluding three, most subjects had some experience with a mouse and windows. Thirty-five percent of the volunteers were female. About half the subjects had previous exposure to hypertext systems.

One subject withdrew during the first part of the experiment after taking offense with the subject of the Holocaust database, yet most other subjects found the process a very educational and enlightening experience.

5.4. Experimental Procedures
Subjects were given an overview of Hyperties and the kind of tasks that they would be requested to perform. Each subject was first asked
The Christian Social Party channeled this social protest into a full-fledged attack on liberal political hegemony in Vienna. Promising those groups which hated the Jews to put an end to so-called Jewish "power," and downplaying such antisemitism among those who found it distasteful, Lueger dazzled the Viennese bourgeoisie with his oratory and led the Christian Social Party.

Social leaders like Ignaz Seipel, Engelbert Dollfuss, and Kurt von Schuschnigg served as prime ministers of Austria.

Engelbert Dollfuss (1892-1934) headed the last democratically elected government of Austria's First Republic. A politician with an impressive record of government service, Dollfuss served his political apprenticeship under Ignaz Seipel whom he succeeded first as the leader of the Christian Social Party and then as chancellor of Austria (1932-34). Conservative and pro-Catholic in his views, Dollfuss attempted to suppress both the Austrian Social Democratic Party and the Austrian Nazi Party during his tenure as chancellor. Furthermore, he staunchly opposed the idea of Anschluss (union) between Austria and Germany. Despite his antisemitism, which he made no effort to conceal, Austrian Jews regarded Dollfuss as a "bulwark" against Nazism, and Nazis portrayed Jews as steadfast supporters of the chancellor.

Dollfuss served as chancellor during an era characterized by violence and social unrest.

A few Christian Socialists became leaders of Austria. When did they first enter office?

Answers:
A. 1922, 1932, 1934
B. 1926, 1932, 1934
C. 1922, 1929, 1932
D. 1922, 1926, 1929
to answer the questions using the single window version which was identical for both experimental groups. In the second part the subjects were randomly assigned to one of the two treatment levels. In each part subjects were given six practice questions using the recipe database followed by six questions using the Holocaust database. Out of the six questions, three were considered easy and required traversal of one or two articles and three were considered hard and involved a traversal of three or more articles. To reduce bias attributed to question difficulty, the first ten subjects were given the easy questions first. The rest of the subjects were given the questions in reverse order of difficulty.

The practice session included a demonstration and subjects were also advised of the kind of strategies that might help them answer the questions.

5.5. Task

The task was to answer multiple choice questions that were displayed on the bottom of the screen below the four windows. After answering each question subjects had to select the NEXT QUESTION button upon which the software saved the collected information.

Table 1 shows the optimal path for each of the tasks. The numbers represent the minimal number of pages that were needed within each article to find the next link or to find the answer.

<table>
<thead>
<tr>
<th>Question #</th>
<th>Single Window</th>
<th>Multi Window</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>3 1 1 1</td>
<td>2 1 1 1</td>
</tr>
<tr>
<td>5</td>
<td>2 2 1 5 1</td>
<td>1 1 2 3</td>
</tr>
<tr>
<td>6</td>
<td>4 3 4 1</td>
<td>6 5 5</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>35</td>
</tr>
</tbody>
</table>

Table 1: Experimental task breakdown analysis for optimal paths.

5.6. Concerns

The time measurement was inclusive of all task components. Each time measurement was started upon the selection of the NEXT QUESTION button to the following selection of the same button to start the next task. The reason for not recording the time immediately upon answer selection was to allow subjects to correct their answers if they made a mistake or wanted to verify their response.

Only a fraction of the time was spent in the task domain. However, it is almost impossible to isolate the time used for housekeeping from the task domain time, as these processes are internal to each subject's mind.

A software error led to a loss of measurements for the number of articles, backtracking steps and index usage for the last (sixth) task of each experimental treatment. While this reduces the total number of data points available for the analysis, the fact that the order of the questions was reversed halfway through the experiment spread the problem evenly among the measurements for both levels of task difficulty.

5.7. Results

5.7.1. Performance Measurements

All performance measurements were analyzed using a two-way ANCOVA with a repeated measure in the task difficulty treatment. The two control strategies were considered the main treatment, and are referred to as treatment A. The level of task difficulty, was considered the secondary treatment and is referred to as treatment B.

Time was measured as the mean time per correct response per subject per difficulty level. Results for speed measurements showed that subjects completed tasks in the Automatic Panning faster than for the User Controlled Placement, though the results were not statistically significant at the p<0.05 level (F(1,17)=2.1). On the second treatment, task difficulty, the results show statistically significant differences at the p<0.001 level (F(1,18)=47.5). No interaction effect was found.
Results for accuracy did not show significant differences in any of the treatments or their interaction due to a very low error rate, although the number of errors for the harder tasks was slightly larger.

The number of articles in a forward traversal per subject per task difficulty was measured. The results show statistically significant differences at the p<0.001 level (F(1,18)=59.1) that the more difficult tasks required longer traversal path to correctly complete the problem. This result affirms the classification of the tasks by difficulty.

The UNDO feature allowed subjects to backtrack, one step at a time, to previously retrieved articles. This feature was extremely helpful to subjects in the Automatic Panning when a single article presented a branching choice for the candidate retrieval. If the retrieval led to a dead end, or if multiple comparisons were needed in relation to the branching point, the UNDO allowed a quick return to that branching point.

Backtracking was measured as the number of UNDO calls per subject per difficulty level. The results indicate statistically significant differences at the p<0.01 level that the number of backtracking was directly correlated to the task difficulty. Also, the number of times the UNDO was used with the Automatic Panning was significantly higher than with the User Controlled Placement. There was also a statistically significant difference for the interaction between the treatments at the p<0.05 level. In addition, the heterogeneity of the regression test was found to be significant, too, indicating that the significance in the main effects may have been contributed to by other factors which were not accounted for by the experiment.

5.7.2. Subjective Evaluation
Participants were asked to rank, on a scale of one to ten (ten is best), the two strategies with which they performed according to: a) ease of use of the particular method; b) how helpful was the method they used to accomplish the task; and c)
how fast did they think they performed. Since all subjects ranked the Single Window control strategy, a one-way analysis of covariance was performed for each criteria.

Analysis of the subjective questionnaire yielded statistically significant results for all three questions at p<0.01 level favoring the User Controlled Placement. The F values in the corresponding order were F=9.0, F=14.8, and F=9.7. Interestingly, though, the heterogeneity of the regression test for the speed ranking was found to be significant at the p<0.05 level. This test may indicate that other factors which were not accounted for by the experiment may have contributed to the obtained result.

The participants were very receptive and provided written comments at the end of the evaluation form. One class of comments referred to the essence of the visual difference, namely that the single window did not provide enough information.

Five of the eleven subjects that were assigned to the Automatic Panning group made comments like “For the most part [I] liked multiple windows, but [it was] a little confusing,” and “The multiple windows, although providing many pages of information, at one particular time, overwhelmed this viewer and provided additional distraction.” Yet, in at least two other cases the subjects noted that they did not utilize the background windows and diverted their behavior to a Single Window strategy; “multiple windows - more confusing - chose to use only one window”. On the other hand, in three other cases the subjects had more positive remarks like “Multi windows...[were] convenient to leave some info on [the] screen,” and “I liked the amount of information displayed by the multiple windows. [It was] easier to retain information.”

Five of eleven assignees to the User Controlled Placement liked the increased flexibility and the ability to place the articles at any window. Their comments included: “I liked the flexibility of getting all the information on the screen at the same time,” and “[I] liked placement choice in multiple window system”. Two subjects noted that they needed to memorize less information: “This way I didn’t need to rely on memory” and “the additional

<table>
<thead>
<tr>
<th>Rank</th>
<th>AP</th>
<th>UCP</th>
<th>AP</th>
<th>UCP</th>
<th>AP</th>
<th>UCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>7.3(2.1)</td>
<td>9.5(0.7)</td>
<td>9.5(0.5)</td>
<td>9.5</td>
<td>8.9(0.8)</td>
<td>8.8</td>
</tr>
<tr>
<td>5</td>
<td>7.3</td>
<td>6.8(2.1)</td>
<td>7.4(1.2)</td>
<td>7.5</td>
<td>8.9</td>
<td>8.8</td>
</tr>
</tbody>
</table>

Figure 4: Results summary for subjective evaluation
windows were helpful in retaining data needed to make final decisions”. Interestingly, two subjects complained that the User Controlled Placement was “just tiring after a while.”

6. Discussion

Clearly, the User Controlled Placement was the favored strategy with regard to ease of use, and more importantly, the ability to support the kind of tasks that were posed to the participants. However, the performance analysis indicates that users did not perform as fast as they thought they did.

All performance measurements, except the one for error rates, clearly show that the task difficulty produced statistically significant differences. These results were expected, and affirm the difference among the difficulty levels of the assigned tasks. The failure to find significant differences in error rates can be attributed to the fact that at the beginning of the experiment it was stressed that accuracy is of high importance. Also the lack of time constraints did not produce the pressure that could have led to increased error rates.

The average number of articles used in the User Controlled Placement was slightly higher. This was a little surprising, especially because there were not many errors found. Several explanations come to mind. From the experimenter observations, users who adjusted quickly to the User Controlled Placement strategy tended to start the search by first filling many windows with candidate articles for their search. This allowed them to batch their motor operations at the beginning of the search and saved them from interruptions later on. However, in many cases they did not need all those articles, and eventually they did not read them after they found the correct answers. Another possible explanation is the lack of concentration due to increased activity outside the task domain.

Surprisingly, index usage was very similar in both browsing strategies. Subjects in the User Controlled Placement group did not take advantage of the fact that all the windows were active, and could have devoted a specific window for all index retrievals. Instead, the index was requested every time it was relevant to search it.

UNDO usage was consistent with the expectations that it would be larger for Automatic Panning because the UNDO plays a major role in using this strategy. The UNDO in the Automatic Panning is the only flexible feature that allows a quick reversibility of retrieval actions.

The subjective evaluation questionnaire provided some insight into the users’ preferences. The reverse tendencies between the actual measured speed and the speed in which users thought they performed are quite striking. One possible explanation is that the participants were biased in evaluating the strategies, believing that having four windows under their control simultaneously must have speeded their performance. Another explanation of this phenomena is that the participants in answering this question considered their time in the task domain only, in contrast to the total time measurements that were recorded by the software.

In conclusion, it is our opinion that User Controlled Placement showed the potential of improving the utilization of hypertext searches when faced with well defined tasks. For simple tasks it seems that a single window or multiple windows with Automatic Panning will be effective and adequate. As a result of our research we believe that, with increased hypertext size and task complexity, the navigational strategy would benefit from greater flexibility such as provided by User Controlled Placement.

7. Acknowledgment

The authors would like to thank Dr. Kent Norman for his assistance and guidance with the experiment design and the statistical analysis and Rick Chimera for his thoughtful comments.
References


