USER INTERFACE DESIGN AND EVALUATION
FOR AN ELECTRONIC ENCYCLOPEDIA

BEN SHNEIDERMAN
Department of Computer Science, University of Maryland, College Park, MD 20742

ABSTRACT
The Interactive Encyclopedia System (TIES) has been under development since Fall 1983. It enables users to easily traverse a database of articles by merely pointing at highlighted words in context. This embedded menus approach to hypertext and its user interface design are described with three exploratory studies of TIES use. Plans for future development and studies are offered.

1. The Dream

Printed books were an enormous stimulus to science, culture, commerce, and entertainment. Electronic books may produce a similar stimulus in the next century, but current designs are poor. Typical screens are too small, too slow, too complicated, and too hard to read. With careful attention to the user interface and the underlying technology, we have a chance to create a new medium that is potentially more attractive and effective than printed books in many situations (Yankelovich, Meyrowitz, and Van Dam, 1985; Conklin, 1986).

Electronic books can have color, animation, sound, rapid access, compactness, rapid traversal and search, user annotation, electronic dissemination and updating, dynamic text to reflect the user’s needs, and other features yet undreamed of.

Our project is called TIES (The Interactive Encyclopedia System) to describe the overall concept and to draw attention to the way in which articles are conveniently linked (tied) together (Shneiderman and Morariu, 1986). However, TIES has taken on several other connotations:

The Interactive Exploration System
The Interactive Exhibit System
The Information Exploration System
The Interactive Explanation System
The Information Expert System

TIES can be all these and more. It can be an addition to a museum exhibit, a browser for organizational policy manuals, a tool for diagnostic problem solving, an environment for novels or mysteries, an online help strategy, a browser for computer program text, a format for cookbooks or self-help manuals, or a way to explore cross referenced materials such as legal documents or an annotated Bible.

2. The Past

The Interactive Encyclopedia System (TIES) has been under development at the University of Maryland since Fall 1983. It allows novice users to explore information resources in an easy and appealing manner. They merely touch (or use arrow keys to move a light bar onto) topics that interest them and a brief definition appears at the bottom of the screen. The users may continue reading or ask for details about the selected topic. An article about a topic may be one or more screens long. As users traverse articles, TIES keeps the path and allows easy reversal, building confidence and a sense of control. Users can also select articles from an index.

TIES is attractive for instruction (and entertainment) because the author’s ideas and writing style are the focus of attention. Through careful human factors design, the computer aspects have been trimmed to let the author communicate to the students and to allow the students to control their learning. TIES has a defined instructional strategy - articles linked by embedded menus - that simplifies design, permits strong support for authoring, and enables use of previously written materials.

TIES consists of:

- Novice user browsing software
- Authoring software for composing new articles and editing
- Databases for browsing.
Databases already created:
Austria and the Holocaust (100 articles)
Adele Stamp Student Union (42 articles)
Introduction to TIES (18 articles)
Introduction to Authoring on TIES (16 articles)
FULCRUM Project (30 articles)
Online Maintenance Manual (52 articles)
Managing your Credit (23 articles)
David Seymour (1911-1956): Photographer (30 articles)
Wines (41 articles)
UM Office of Minority Student Education (41 articles)

TIES is appealing to authors because of the explicit instructional model, the reduction of computer-related concepts, the focus on content, and the lively user interface. It is an engaging challenge to reformulate pedagogic plans into the network of related articles that TIES supports. There is a great sense of satisfaction in composing articles and seeing the linkages come to life as they are used by students in novel ways.

TIES allows authors to create a network of conceptual knowledge in which concepts are linked explicitly and the readers are free to explore pathways based on their needs and interests. Potential applications include tutorials about software products, self-study guides for any discipline, travel guidebooks, annotated Shakespeare or the Bible, and maintenance manuals for equipment. Each visitor suggests intriguing and novel applications.

The TIES authoring software guides the author in writing a title, brief definition (5-35 words), text (50-1000 words, typically), and synonyms for each article title. Authors mark references in the text by surrounding them with a pair of tildes. TIES collects all references, prompts the user for synonym relationships, maintains a list of articles, and allows editing, addition, and deletion of articles. The author tool displays TO/FROM citations for each article and allows authors to keep notes on each article. A simple word processor is embedded in the authoring software, but users can create articles on their own word processor, if they wish. There are no commands to memorize, every operation is done by selection from options on the screen.
TIES runs on a standard PC (monochrome or color) and on PCs, XTs, or ATs with or without touchscreens. We are attracted to the possibility of eliminating the keyboard while still providing substantial exploratory power. TIES was first written in APL and has been rewritten in the C programming language. Dan Ostroff, a graduate student in computer science, did the implementation and a major portion of the user interface design. Dr. Janis Morariu of the Center for Instructional Development and Evaluation contributed substantially to the user interface design. Jacob Lifshitz and Susan Flynn maintained and enhanced the system during 1985-86.

3. Early Empirical Studies

Several experimental studies have been conducted to test out certain design alternatives (such as demonstrating the advantage of arrow keys over the mouse for this system) and observe user behavior. More than 220 subjects participated in these controlled experiments. In addition, more than three hundred novices and experts have tried and commented informally on the current design.

In the study comparing the arrow keys (maybe better termed "jump" keys because the cursor would jump to the closest target in the direction pressed) to the mouse, the arrow-jump keys proved to be an average of 15% faster and preferred by almost 90% of the subjects (Ewing et al., 1986). We conjecture that when there are a small number of targets on the screen and when arrow-jump keys can be implemented, they provide a rapid, predictable, and appealing mechanism for selection.

The TIES approach of embedded menus (Koved and Shneiderman, 1986) idea was also used for two experiments with online maintenance manuals for electronic equipment. A tree structured and linear form of a 52-page maintenance manual was prepared for screen presentation and in paper form (Koved, 1985). Experimental subjects had to perform 12 tasks using one of the manuals. Significant differences were found showing that time was reduced using the paper versions. No significant differences were found between the tree and linear versions for speed or error rates. When a pruning algorithm was applied to the text to trim text unrelated to the task, the time was cut in half. This latter experiment used only the computer condition and demonstrated one of the advantages of screens over printed text. This is important, since for many applications
printed manuals are still easier to use and approximately 30% faster to read than computer displays.

Four selection strategies (mouse, arrow-jump, touchscreen, and numbered keys) were compared in a study in which each of 24 subjects used each strategy to follow prescribed paths through the database (Ostroff, 1986). The touchscreen resulted in the fastest time and highest preference, but also the highest error rate. Performance and preference for the arrow-jump strategy was close to the touchscreen. Improved strategies for touchscreen and the mouse were proposed.

Field studies have also been productive in gaining feedback from users in realistic settings. A week long trial was held during 1985 in a museum in Washington, DC with careful observation of each user and generous amounts of personal attention when needed. On May 21, 1986 a version of TIES was installed for unattended use in downtown Washington, DC in conjunction with an exhibit of the photographs of David Seymour. On November 14, 1986 a similar version of TIES was installed in the International Center of Photography at 94th Street and Fifth Avenue in New York City with another exhibit of David Seymour’s photos. We are collecting usage data to understand reading patterns for patrons of these exhibits.

4. New Research Results

This paper presents three exploratory studies in which novice TIES users had to traverse a database to answer questions posed by the experimenter. This mode of usage is mid-way between the required paths of the early studies and the informal browsing during the field studies and museum installations.

4.1 Embedded vs. Explicit Menus

This exploratory study, conducted by David Powell, compared embedded menus with the traditional approach (as exemplified by pioneering systems such as ZOG (Robertson, McCracken, and Newell, 1981)) of text followed by a menu of choices for further information. The traditional approach is easy to implement and neatly separates the reading
of informative text from the selection of further information. On the other hand, the embedded menus approach seems attractive because choices are shown in context and users can select them immediately rather than paging to the end of the text.

Materials. A database of 42 brief articles (1-3 screens) about the services of the University of Maryland's Adele H. Stamp Student Union was prepared in both formats. The database had a strictly hierarchical design with matching halves in similar structures and complexity. Two sets (one for each half) of twenty short-answer questions were prepared for the performance tasks. A subjective questionnaire was also developed.

In the embedded menus version subjects could select highlighted words immediately and view the selected article (Figure 1). With the explicit menu version subjects saw similarly highlighted words, but they had to turn one or occasionally more screens till they reached the explicit menu (Figure 2). Then they typed a number and pressed enter. Even if a word appeared highlighted several times in the text it appeared only once in the explicit menu. The items in the explicit menu were in the same order as their first occurrence in the text. The INDEX command was eliminated during this study.

An IPM PC and a monochrome display with 24 lines by 80 characters was used in this study.
The Adele H. Stamp Union, formerly known as the Student Union, is the cultural and social center for the University. The Union provides a variety of services to the faculty, staff, and students. A plethora of restaurants are available providing a wide choice of atmosphere and a variety of menus. The Union is also a center for entertainment. Several shops and many special services are available, too. Union programs include concerts, exhibitions, and craft classes. The Union is open all week from 7 AM to 1 AM, Monday through Friday, and until 2 AM on weekends.

[Select option then press RETURN]
INTRODUCTION: STAMP UNION

Would you like to read more about
1. RESTAURANTS
2. ENTERTAINMENT
3. SHOPS
4. SPECIAL SERVICES
5. UNION PROGRAMS
6. Return to first page of STAMP UNION

Please enter one number and press RETURN:
Enter Q to quit the session

Figure 2: The explicit menu version separates the reading of text from the selection of articles. Note that the items in this menu are the highlighted items from Figure 1.

Subjects. Twenty-one females and fourteen male undergraduates, mostly first and second year students from the Psychology Department subject pool, participated in this study. The average age was just under 20 and the standard deviation was 1.55.

Procedures. Each subject worked with the embedded menu and the explicit menu version switching question sets when they switched versions. The experiment was an unorthodox design that exposed subjects to both versions and both question sets, but prevented use of standard within-subjects statistics. For simplicity, results with embedded or explicit menus were grouped, independent of order or question set. The study was conducted one subject at a time in a quiet and private room. Subjects signed a consent agreement, completed a brief pre-test background questionnaire, and were given a five minute practice session with a different database. Then they were tested for 15 minutes on each version, and completed a post-test questionnaire.

The tasks required simple short answers such as what hours a restaurant was open or which room the director was located in. Questions had to be answered in order on paper, but the computer collected time, keystrokes, and other data. One point was
given for each correct answer. The few wrong answers were simple slips such as substitutions of a phone number for a room number or misspellings.

**Results.** When using the embedded menus version subjects correctly completed an average of 15.2 questions, but only 12.2 questions with the explicit menus version ($p < .001$). With the embedded menus, subjects viewed an average of 44.2 screens in 31.9 articles, while with the explicit menus subjects viewed 54.32 screens in 17.9 articles (both differences were significant at the $p < .001$ level). These differences indicate that with embedded menus subjects frequently jumped among articles (seeing an average of about one and one half screens per article), while with explicit menus subjects remained within an article longer (seeing an average of about three screens per article) in part because they often had to get to the explicit menus at the end of the article.

The post-test subjective question asked: Do you like using this version of TIES? where 1 was dislike and 7 was like. The average over all 35 subjects was 5.9 for the embedded menus version and 4.2 for the explicit menus version ($p < .001$).

**Discussion.** The results indicate about 25% improved performance speed in answering simple queries using the embedded vs. the explicit menus. Subjective preference was strongly in favor of the embedded menus version. Apparently subjects found it rapid and convenient to recognize a vital highlighted word and to immediately jump to the related article, as opposed to turning screens until they reached the explicit menus at the end of articles. One subject commented "[embedded menu] was a little easier to manipulate -- I think it flowed more smoothly than the [explicit]." Another noted that "I find the [explicit] menu step kind of a bother..." However some subjects did appreciate the explicit menu version "because you can go to the list and pick what you're looking for without having to search through the paragraph."

Since the questions sought specific information and they were presented to the subjects by the experimenter, it might be argued that the explicit menu approach was favored because subjects could often rapidly scan the compact explicit menu. It would be interesting to try a similar study with exploratory browsing by discretionary users. A further issue is that the explicit menus were presented in the traditional manner at the end of the articles. It would be interesting to see how the results would change if the explicit menus were presented on the first screen of an article.
In summary, performance and subjective satisfaction scores suggest an advantage for the embedded menus approach, but replication with more careful experimental methods, with various tasks, and with alternate layouts would be useful.

4.2 TIES vs. Paper

A natural question to ask is whether TIES is as effective as paper in scanning large textual databases. Many studies have shown that reading continuous text or proofreading from screens is approximately thirty percent slower than working from typewritten paper documents (Mills and Weldon, 1986; Gould and Grischkowsky, 1984; Shneiderman, 1987). When additional delays for entering page turning commands and system response time are considered computerized versions can take twice as long as paper versions. Higher resolution screens can reduce the penalty for the reading component. Simpler commands, faster display rates, and faster response times are also helpful in reducing the difference. In fact, there are situations in which computer-based systems would be expected to be faster than paper versions (Koved, 1985).

Jadin Sukri conducted an exploratory study to compare question answering performance by novices using TIES and a paper version of the database. Based on previous studies we expected that paper versions would be faster than TIES, but we wanted to ascertain the magnitude and to study TIES usage patterns.

Materials. The database on "The Holocaust and its Origin: The Case of Austria" (written by Dr. Marsha Rozenblit and Darla Courtney) containing 106 articles (50 to 1000 words each) was used. The paper version was printed out single spaced with 50 lines per page and 60 characters per line. Pages were numbered and a table of contents was prepared. Each article began on a new page, producing a single-sided bound book of 138 pages.

TIES Version 2.1 was used on an IBM PC/AT with a hard disk. The Enhanced Graphics Adapter was used in 43-line mode, yielding 18 lines of double-spaced text. The same database was used for the online version.

Subjects. Twenty psychology undergraduates from the University of Maryland
participated in this study, receiving extra credit for their psychology courses. There were 15 females and 5 males, with age ranges from 18 to 23 years. Eleven had never taken a computer programming course and 15 claimed to never or rarely use computers.

**Procedures.** Two sets of questions (A and B) were prepared and one-quarter of the subjects worked on TIES with A, TIES with B, Paper with A, and Paper with B. Then for the second set of questions one quarter of the subjects worked on Paper with B, Paper with A, TIES with B, and TIES with A, respectively.

Each set of questions had three tasks:

Task 1: simple fact at start of article

Task 2: fact in body of article

Task 3: answer required search in two articles

Before the exploratory study, subjects filled out a brief background questionnaire and signed a consent form. Subjects were given a practice task for 5-8 minutes on either the TIES or Paper versions and then worked on the questions without assistance. Answers had to be written on a paper form. Timing for TIES was done by the computer; timing for the Paper versions were done by stop watch. After the study subjects responded to a subjective satisfaction questionnaire about their experiences with both strategies.

**Results.** Subjects worked through the questions easily in both forms. The timing data indicated that paper was statistically significantly faster than TIES for the simple tasks, but that for Tasks 2 and 3 there was no significant difference (see Table 1).
<table>
<thead>
<tr>
<th></th>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper</td>
<td>21.7</td>
<td>50.7</td>
<td>106.9</td>
</tr>
<tr>
<td></td>
<td>(5.2)</td>
<td>(14.0)</td>
<td>(21.9)</td>
</tr>
<tr>
<td>TIES</td>
<td>41.7</td>
<td>57.8</td>
<td>106.8</td>
</tr>
<tr>
<td></td>
<td>(8.2)</td>
<td>(16.0)</td>
<td>(21.5)</td>
</tr>
<tr>
<td>p</td>
<td>&lt;.001</td>
<td>.148</td>
<td>.986</td>
</tr>
</tbody>
</table>

Table 1: Mean time in seconds (standard deviations) for tasks in Paper and TIES versions. Significance levels by repeated measures t-tests are in the last row.

The post-test questionnaire revealed a positive attitude toward TIES. Sixteen expressed a preference for using TIES, two for paper, and two neutral. Even though the time was longer with TIES or equivalent, eight subjects perceived that TIES was faster, while only four thought paper was faster. Thirteen found paper more tiring, while only two found TIES more tiring. All twenty selected "interesting" as opposed to "boring" or "tiring" when questioned about TIES.

**Discussion.** This exploratory study helps refine our understanding of the utility of TIES for serious question answering from a substantial textual database. We were pleased to find that TIES performed well in time measures compared to paper and that subjective preference was in favor of TIES. This supports our impression that the embedded menus approach is attractive and easy to use for novices. Subjects expressed confidence in their ability to traverse the database and did not feel lost.
Our next steps are to understand better where TIES users were having difficulties and to examine still more complex tasks. Another direction would be to determine how users decided when to use the index and when to select articles. We would also like to know how different database organizations and article writing strategies could improve the utility of TIES databases.

4.3 Screen size (9, 18, and 34 lines)

One of the issues that we have dealt with regularly in the design of TIES is question of screen readability and information density. Early versions used all 80 characters on each line and 18-20 lines of text. This often felt very crowded even to frequent users and therefore we moved to using double spacing and six character margins on the right and left (68 characters of centered text). This produced a 9-line display of text (612 character maximum compared to a full screen of 1920 characters) that seemed appropriate for the museum installations where novice and potentially anxious users predominate. Our informal design decision was based on intuition and reports in the literature (e.g., Gould and Grischkowsky, 1983 or Shneiderman, 1987) that indicated that screen readability was a problem. Later versions of TIES permitted authors to select margin size and line spacing.

In an attempt to investigate screen readability and performance, Steve Versteeg conducted an exploratory study that varied the number of lines on the screen. To control for multiple possible biases, he used only single spacing but varied the number of lines from 9 to 18 to 34 using the Enhanced Graphics Adapter that shows up to 43 lines. The line size was maintained at 68 characters. The smaller fonts were perceived as more difficult to read and so there was concern that the 34 line version would be a challenge. On the other hand more lines per screen would reduce the need for page turning.

A recent study by Reisel and Shneiderman (1987) had shown that increasing screen size from 10 to 22 to 60 and to 120 lines on an IBM 3200 Plasma Display resulted in improved performance and preference for program reading.

Materials. The same database on "The Holocaust and its Origins: The Case of Austria" from the previous study was used but a different set of questions was prepared.
Tasks included multiple choice questions such as "Who formed the Christian Social Party?" or "When did American troops liberate Mathausen?" Subjects answered five questions with each version. A post-test subjective satisfaction questionnaire probed the subjects' preferences.

The single-spaced display of 9, 18, or 34 lines of 68 characters began near the top and the bottom part of the screen was left blank for the first two treatments. The Enhanced Graphics Adapter and display were connected to an IBM PC/AT with a hard disk.

Subjects. Twenty-one subjects from the University of Maryland Psychology Department subject pool participated in the study, but data on only twelve subjects was used for the statistical analysis. Administrative failures eliminated four subjects and five subjects were randomly eliminated.

Procedures. The subjects were randomly assigned to different orders of presentation in a Latin square design. Subjects signed a consent form and were shown how to move the cursor and make selections. Questions were presented on paper but the answers were captured by the computer. Subjects had to press the F1 key and indicate the letter of their choice. This enabled accurate time measurement. Subjects were told to work so that they did not make errors, but that timing data was being recorded. After the three treatments subjects filled out the questionnaire and were debriefed.

Results. The mean times for completion using the 9-, 18-, and 34-line versions were 634, 593, and 513 seconds, respectively, but the result was far from being statistically significant. The ranges for the twelve subjects in each treatment were (446, 1625) for the 9-line, (218, 716) for the 18-line, and (221, 948) for the 34-line versions.

The mean number of page flips ranged from 25 to 16 to 3, demonstrating the advantage of a more densely packed screen (p < 34-line version. Nine preferred the 18-line version and three preferred the 9-line version (p < .001 by chi-squared test).

A statistically significant order effect was found indicating that the first trial took longer than the second or third trials independent of version. An analysis for just the second and third trials was done, but still the time differences among versions were not
During the interview, subjects generally commented that the task was "fun." They added that after they overcame their confusion during the first treatment, they found the questions easy to answer and that they did not "get lost" in the database.

**Discussion.** The small number of subjects and large variability in performance led to a failure to demonstrate statistically significant differences in performance times. However the reduced mean times for increased numbers of lines per screen suggests the possibility of benefit. However, the crowded screen and small font certainly made the densely filled screen appear to be difficult to read. Several subjects in the 34-line treatment placed their fingers on the screen to help them keep their place while reading. These difficulties resulted in a preference for the 9- and 18-line versions. It is possible that with more experience subjects would lose their anxiety and become comfortable with the higher density.

These results make us cautious about the use of the 34-line version in museum applications, but encourage us in using it for frequent users. Our preference is to find screens that can display more lines of text while providing adequate interline spacing. We will also seek screens that have higher resolution and more readable fonts. Our next step experimentally is to explore different margin sizes and interline spacing with novice and more experienced users.

**5. Future Directions**

In June 1986 we completed TIES 2.1 which contains support for display of graphic images. Version 2.2, completed in March 1987, revised the formatting and editing routines and made some minor changes to the user interface. Current implementation efforts focus on support for videodisc and improved index usage. Further testing is needed to develop strategies for touchscreens or other selection devices. We are working with the IBM InfoWindow system and the Microtouch screen attached to an IBM PC/AT. Simultaneously we are testing TIES and running experiments on the AT&T 6300 Plus personal computer.
An advanced browser with string search, various bookmarks, multiple windows, user annotation, printing, etc. is being designed. An advanced authoring tool with automatic reference marking and sophisticated editing is also planned.

Jacob Lifshitz has implemented TIES on larger screens (SUN3 workstations) to allow multiple large windows and eventually pointing at graphics objects. We are working to make TIES a useful exploratory tool for expert users. At the same time we hope the personal computer versions with touchscreen will serve the large fraction of the population that is still anxious about using computers.

The University of Maryland has contracted with a software development firm to support commercial applications and distribution of TIES: Charles Kreitzberg, President ( (909) 799-5005 ), Cognetics Corporation, 55 Princeton-Hightstown Road, Princeton Junction, NJ 08550.

References


