

An empirical comparison of the usability for novice and expert searchers of a textual and a graphic interface to an art-resource database

Andrew Dillon and Min Song
School of Library and Information Science
Indiana University
Bloomington IN 47405
email adillon@indiana.edu, msong@indiana.edu

This item is not the definitive copy. Please use the following citation when referencing this material: Dillon, A. and Song M. (1997) An empirical comparison of the usability for novice and expert searchers of a textual and a graphic interface to an art-resource database, *Journal of Digital Information*, 1(1).

Abstract

The present paper reports an experimental test of a prototype graphic and textual search interface for a university database on art-resource works. Novice and expert searchers were tested on both interfaces with performance assessed in terms of search speed and accuracy. Verbal protocols and navigation strategies were also examined. Experts performed significantly faster than novices though both user groups performed slightly (but not significantly) faster with the graphical interface. Furthermore, the graphical interface significantly reduced navigation effort. While there were no significant task accuracy differences, novices failed to complete more searches with the textual interface. Implications of these results for search interfaces to digital resources are briefly discussed.

1. Introduction

Information retrieval (IR) is generally considered a basic information literacy requirement in the digital age though it is widely recognized that novice users often perform poorly with existing IR tools (see e.g., Allen, 1994). Frequently, novices experience problems formulating valid queries while related issues of poor interface design on IR systems serve to complicate the search process.

The surge in use and popularity of graphical user interfaces (GUIs) over the last 10 years has been attributed to both technological enhancement (the availability of higher quality monitors which render it possible to produce distinct, detailed images on the screen) and theoretical assumptions e.g., a suitable graphical representation is easier to comprehend (Lodding, 1983) and direct manipulation affords a greater sense of user control (Shneiderman 1992). Specifically, experiments on human memory for pictures or the effects of imagery on memory (e.g., Mandl and Levin, 1989) lend support to the broad notion of iconic representation in graphical interfaces enhancing usability compared to text-based interfaces. Similarly the control of interfaces afforded by a mouse or equivalent input device is considered to be optimal with respect to human information processing (Card et al, 1983).

Despite the increased use of GUIs in IR applications and Web search engine interfaces, it is not clear that such interface features enhance usability in all situations. With respect to icons for example, Lansdale and et al. (1989 and 1990) showed that any gains in speed of the users visual scanning process could be at the cost of a greater likelihood of missing the target. Other research suggests that it is less the iconic representation than the associated use of text labels on icons that aids users in GUIs (for a good review of such work see Rogers 1989).

One area where icons in combination with direct manipulation may offer significant advantages over text input is in terms of selecting objects on a screen. Here, established models of human performance in target selection (e.g., Card et al's (1983) Model Human Processor and GOMS analysis framework) could be used to demonstrate quantitatively any speed or efficiency advantage. However, in IR applications, searching is less a matter of hitting targets on a screen than formulating queries that are conceptually appropriate, and in such cases it is not at all appropriate to recast the task in terms of target selection

and mouse movement. As a result, it is not logically the case that the use of GUIs will enhance searcher performance.

Drawing up reliable and valid recommendations for interface features for any task is complicated further by the possible interaction of user type on performance with particular interface styles. Chui and Dillon (1997) report data showing that users experienced with animation in graphical interfaces are more likely to be affected by its presence or absence than users with no experience of animation. In information retrieval tasks, it has long been known that there exist large individual differences in performance (see e.g., Fenichel, 1981) and it is possible that visual components of search interfaces might be an important source of this variance, or may effect novice user performance with an IR system. Allen (1994) demonstrated an interaction between logical reasoning ability and interface style in an information retrieval application, and Sein et al (1993) showed that users' scores on a test of visual ability served as a predictor of learning success on a graphical interface. Thus, individual differences among users seem to be a significant source of variance in determining the usability of all IR interface components.

In the present paper we report a usability evaluation of a prototype IR interface aimed ultimately at supporting novice user access to library materials on art. While the study fed into the early stages of the design process of a new tool, we were mainly interested in seeing how two key properties of a graphical interface, (the use of icons and the capability to select options with a mouse) might improve novice searcher performance and perceived satisfaction with the tool.

2. Test System Design

A prototype database containing 250 bibliographic records of picture-related information (art works, books on artists, collections, catalogues etc.) available to users of Indiana University Libraries was developed as a test bench. One artist and one reference librarian at Indiana University, both experts in the field of art and humanities, were involved in designing a graphical interface for this database that would serve as an alternative to the standard text database of the university catalogue (see Figures 1 (text interface) and 2 (experimental interface)).

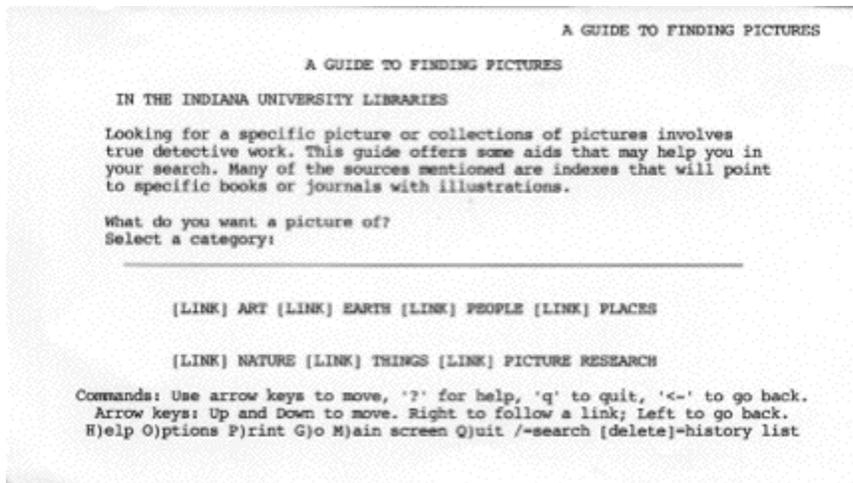


Figure 1 Textual interface

The major differences between the interfaces from the users' point of view were the ability to point and click (enabled in the graphical) instead of typing commands (required in the textual interface), and the presence of images (as opposed to text) supporting the categorisation of the database contents in the graphical interface..

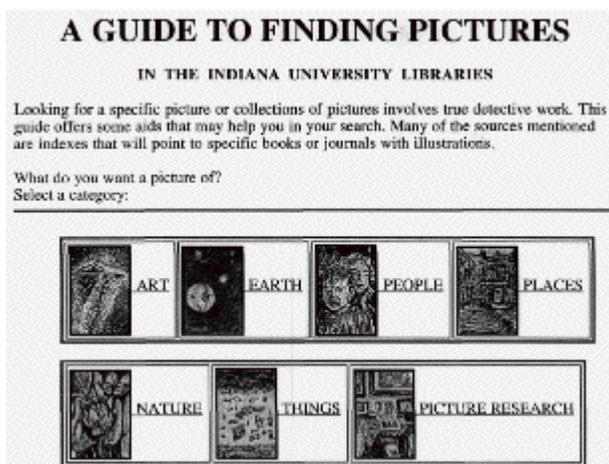


Figure 2 Graphical interface

The data in the system mainly consist of reference materials in the fields of art and art history. The initial design change was aimed at providing two levels of iconic representation above the final output of relevant resources and their locations (which was identical for both interfaces). The data structure adapted for both interfaces was hierarchical. At the root node, the categories for classifying the data are based on broad concepts such as name of artist, picture, and scripture. Each category was broken down to

subcategories. Icons were designed to represent the underlying object or function to which they refer. In order to maximize the effectiveness of graphical representation, pictorial icons were drawn by an artist, one of the design team members. The system does not allow user to do keyword searching but supports browsing only. Apart from the interfaces both databases were identical in content and structure.

3. Research Design

Subjects

Two groups of users (12 experts and 12 novices) participated in the study. The expert group (10 females/2 males, ages: 25-45) consisted of four reference librarians at Indiana University and eight graduate students in IU SLIS (School of Library and Information Science). All students were taking or had taken the graduate level Online Searching course at the time that they were recruited and had substantial experience of online searching with a variety of electronic databases of both forms. The novice group consisted of 7 freshmen and 5 sophomore students at Indiana University majoring in Arts and Humanities programs (6 female, 6 male, age range: 18-22). All reported no prior online searching experience and less than one year total computer experience.

It should be noted here that the user type variable manipulated expertise in searching with IR systems, not domain (art) expertise. The domain variable that is likely to be worthy of further investigation but is not central to the focus of this system design.

Experimental design

A two factor design compared all users on both interface styles. Independent variables were user type (novice/expert), and interface style (graphical/textual). Dependent variables were time taken to complete tasks, accuracy of search performance, navigation style (operationalized as number of nodes visited in the database in each search task) and responses to a post-task interview.

Subjects were required to answer a set of 10 questions (see Appendix). These questions were developed by the authors to ensure that the questions did not unduly favor any one interface and that users would need to explore the database fully to find the answers.

Procedure

All subjects were tested individually at the same workstation, located beside the reference desk in the main library, the intended location of the finished system. Before the search session, brief instructions and training were given by one of the experimenters. At this session, the experimenter described the nature of the investigation and introduced the subjects to the systems and tasks for 15 minutes. The subjects were told that they would be interrupted between questions to allow the experimenter to ask some questions related to the search. Subjects were then given a set of tasks and asked to attempt all questions in the presented order. The subjects were encouraged to verbalize their thoughts. Their comments and movement through screen were recorded by the experimenter.

After five tasks, subjects took short break and then proceeded as before on the second interface. Order of presentation was counterbalanced across all subjects with half the experts and half of the novices starting on the textual interface, and the other half of both user groups starting on the iconic interface. Order of questions was randomized for each subject (there were no a priori reasons for allocating certain questions to specific interface conditions).

4. Results

Results are presented below for each category of data: completion time, accuracy score, navigation, and structured interviews. We used a two-factor ANOVA design (with one repeated factor) to analyze the relevant data of the study.

Completion time

Time taken to complete the 10 tasks was recorded for each subject and a significant main effect was observed for user type ($F_{[1,22]} = 20.73, p < .01$) with the novice group performing significantly slower than experts on both interfaces (see table 1) Obviously search experience is an important influence on speed of task performance, with novices taking significantly longer than experts to use either system (almost twice as long on average).

While no main effect was observed for interface style ($F_{[1,22]} = 1.19, p > .05$), by examining the time data in full (see Table 1), we can see that experts performed approximately 15% faster, and novice users about 12% faster, with the graphical

interface. In usability terms, such differences could be important (see e.g., Landauer, 1991). There was no interaction effect.

<i>User Type</i>	<i>Interface</i>			
	Textual		Graphical	
	Mean	SD	Mean	SD
Expert	421.1	141.0	359.9	131.0
Novice	712.6	314.1	634.3	234.7

Table 1. Mean times and standard deviations per task (in seconds) for each user group.

Accuracy

Since satisfactory answers to many search tasks could take several forms, in the present study accuracy was assessed by awarding two points for an unambiguously correct answer, one point for a partly correct answer (e.g., those that seemed to meet some of the goals of the task but lacked either details or sufficient number of results to be complete) and no points for a wrong answer or an abandoned question. Summary data is presented in Table 3, where maximum score in each category is 10 (5 tasks, maximum 2 points each).

<i>User Type</i>	<i>Interface Styles</i>			
	Textual		Graphical	
	Mean	SD	Mean	SD
Novice	5.4	1.6	5.7	1.9
Expert	6.0	1.4	6.9	1.8

Table 2. Mean and standard deviation for accuracy scores.

Differences between novices and experts approached significance ($F_{[1,22]}=3.07, p<.07$) with experts performing slightly more accurately than novices. A general trend favoring the graphical interface is apparent but is not statistically significant. It should be noted that 4 novice users gave up searching with the textual interface (one each on question 1 and question 6, and two gave up on question 3) and one novice user gave up the search with the graphical interface (question 2). All experts completed all tasks. Scoring

accuracy only on a 3 pt. scale with high rates of non-completion introduces possible ceiling and cellar effects and thus may have served to constrain variance in these data.

Navigation

Navigation was observed by tracing the paths users followed through the database. To gain a general measure of navigation we calculated the number of transitions made between various nodes. Such data can be interpreted in more than one way. Obviously, users who are in difficulty are likely to visit more nodes as they seek to locate relevant information, thus high navigation scores would be a sign of the user having a poor sense of the database's structure. On the other hand, visiting multiple nodes might be seen as a sign of user comfort with exploration, although this interpretation is less common in the hypertext and menu navigation literature (see e.g. Norman, 1991) and our observations of the users in this study support the more conventional interpretation we employed (see section on verbal protocols below).

Analysis indicated a significant effect for interface style ($F_{[1, 22]} = 4.606, p < .05$).

Tukey's HSD follow-up test for interface styles revealed that the number of nodes visited by both user groups were significantly higher with the textual interface than with the graphical interface ($p < .05$). These data are important as they suggest that users were able to gain more direct access to relevant information when icons were employed (see Table 3). For novices this translates approximately into a 15% efficiency gain in navigation, for experts, 10%.

<i>Interface Styles</i>					
		Textual		Graphical	
<i>User Type</i>	Mean	SD	Mean	SD	
Novice	20.17	5.13	17.25	2.98	
Expert	17.75	3.82	16.00	2.62	

Table 3 Nodes visited (mean and standard deviation) for each group

Preference ratings

Participants were asked to express their preference for the interfaces in the search sessions and any major difficulties they experienced. Nine of the 12 novices expressed a

preference for the graphical interfaces, largely based on feelings of control over the interface. Three novices preferred the textual interface, positively commenting on its apparent higher speed. Experts' views were more diverse, with four subjects preferring the graphical interface, three the textual interface, and the remaining five subjects expressing no preference.

General observations from verbal protocols

Verbal protocols were not formally recorded but the experimenter made a note of interesting comments as they occurred and checked interpretations with users after the trial. Generally, novice users tended to spend more time reading the screen and seemed to be more careful about where they searched for answers. Expert users were more decisive. After the search session, most experts commented that time is a critical factor for online searching and they did not want to judge the relevance of records while searching, suggesting a possible contamination of experimental issue with real world knowledge (even though all users were asked to perform the tasks as quickly as possible) Meadow and et al. (1995) reported similar findings in their study where the domain specialists with no search experience tended to spend much more time reading than search specialists.

Both user groups' protocols indicated a tendency to get lost more easily on the textual interface as demonstrated in the navigation measures reported above. Novice users in particular had many difficulties in navigating the textual interface. As noted earlier, four novices gave up answering certain questions.

5. Conclusions

Expert searchers performed well on either interface in the present study. Clearly, advocating graphical interfaces as a global source of improvement for information retrieval systems is not the answer for all users. The significant differences between user types is further confirmation that search expertise is the most important predictor of user performance with an IR system. Use of icons seemed to improve slightly novice performance as measured in this test, primarily in terms navigation through the system and slightly in terms of speed. Novices manifested more efficient search behaviors in the graphical environment in terms of nodes visited to obtain answers and the majority preferred the iconic to the textual interface in this trial.

Obviously the task type may be a major source of the variance in user response to interface style, and searches involving less abstract and more textual material , e.g., title or author searches, may not be so affected by graphical interfaces at the search system. This is one area for further research. But in the present context, the effect on navigation in a comparatively small database is likely to be magnified as the application grows to include many more entries. Thus, the use of the graphical interface is likely to enhance user performance for typical users of the full database in the library.

Acknowledgments

The authors would like to thank IU Libraries for the invitation to be involved in the development process for this database and permission to publish the results of this user study. Comments from both reviewers led to substantial improvements in the quality of the final paper.

Appendix

The Search Tasks

1. Find examples of the printmaker Dore's work.
2. Locate several books with maps of Chicago.
3. Find a US geological map of Mississippi.
4. Find a current map of Asia
5. Find a source for multiple pictures of different butterflies.
6. Find several examples of Impressionist painting?
7. Find one book with illustrations by Picasso.
8. Find examples of Afro-American artists.
9. I need examples of colonial American architecture. What book do I have to refer to?
10. Where can I find pictures by the photographer Eliot Porter?

Bibliography

- Allen, B. (1994) Cognitive abilities and information system usability. *Information Processing and Management*, 30, 177-191.
- Card, S., Moran, T. and Newell, A. (1983) *The Psychology of Human Computer Interaction*. Hillsdale, NJ: LEA.
- Chui, M. and Dillon, A. (1997) Who's Zooming Whom? Attunement at the human-computer interface. *Journal of the American Society for Information Science* (in press)
- Fenichel, C. H. (1981) Online Searching: Measures that Discriminate among Users with Different Types of Experience. *Journal of the American Society for Information Science*, vol. 32 pp. 23-32.
- Landauer, T. (1991) Let's get real: a position paper on the role of cognitive psychology in the design of humanly useful and usable products. In: J. Carroll (ed.) *Designing Interaction*. Cambridge University Press : Cambridge MA. 60-73.
- Lansdale, M. W., Jones M. R., and Jones M. A. (1989) Visual Search in Iconic and Verbal Interfaces. In E. Megaw (ed.) *Contemporary Ergonomics, 1989* , London: Taylor and Francis, pp. 422-429.
- Lansdale, M. W., M. Simpson, and Stroud T. R. M. (1990) A Comparison of Words and Icons as External Memory Aids in an Information Retrieval Task. *Behaviour & Information Technology*, 9,2, pp. 111-131.
- Lodding, K. N. (1983) Iconic Interfacing. *IEEE Computer Graphics and Applications* 24, pp. 11-20.
- Mandl, H. and Levin. J. (eds.) *Knowledge Acquisition from Text and Pictures*, Amsterdam: Elsevier Science Publishers. pp. 3-35 1989.
- Meadow, Charles T., Jiabin Wang, and Weijing Yuan (1995) A Study of User Performance and Attitudes with Information Retrieval Interfaces. *Journal of the American Society for Information Science*, 46 pp. 490-505.
- Norman, K. (1991) *The Psychology of Menu Selection* Norwood NJ: Ablex

Rogers Y. (1989) Icon design for the user interface. *International Review of Ergonomics* 2, 159-164.

Sein, M., Olfmann, L., Bostrom, R. and Davis, S. (1993) Visualization ability as a predictor of user learning success. *International Journal of Man-Machine Studies*, 39(4) 599-620.

Shneiderman, B. (1992) *Designing the User Interface*. Boston: Addison Wesley.