Are Web-based OPACs More Effective Retrieval Systems Than Their Conventional Predecessors? An Experimental Study

Abstract: The World Wide Web (simplified here to “Web”) is well-known for its “point and click” graphical user interface (GUI) and hyperlink search and navigate capabilities. When OPACs are placed in this operational context, users can easily hyperlink from a bibliographic display to related search terms, class marks, or bibliographic records. This hyperlinking capability is not available in most conventional text-based OPACs. As more and more users undertake their searches on Web-based information retrieval systems such as OPACs, it is natural to ask, “Are Web-based OPACs more effective retrieval systems than their conventional predecessors?” This paper presents the findings of an experimental study which compared users’ search performance, assessments of ease of use, and satisfaction with search results after use of a Web OPAC or its conventional counterpart. The primary questions addressed by this research center on the influence of two experimental factors, OPAC search interface style and search task level of difficulty, on the dependent variables: actual search performance, perceptions of ease of use, and user assessments of satisfaction with search results. It was hypothesized that Web OPACs would be assessed as easier to use and that they would outperform conventional OPACs when measured by actual search results and users’ levels of satisfaction with search results. Web OPAC searchers outperformed Text OPAC searchers, but search task level of difficulty is a major determinant of search success. The study also found little association between searchers’ level of satisfaction with results and actual search performance.

1. Introduction

The research study described in this paper investigates (1) the comparative retrieval effectiveness and perceived ease of use of Web OPACs and their conventional counterparts and, (2) the phenomenon of false positives in the context of online library catalog (OPAC) use. Unlike their conventional text-based command and menu-driven predecessors, Web-based, graphical, hypertext OPACs hold the promise of easy search interaction and navigation among related subjects and related items. Facilitated by easy to navigate linkages, and searching that does not require prior knowledge of subject headings or class categories and codes, it might be assumed that searchers would rate these Web OPACs as easier to use, more useful and helpful in the search process, and superior in overall performance. This and related assumptions are examined in this study. The existence and network accessibility of two versions of many OPACs, a Web version and a conventional text version, provided the opportunity to proceed with this investigation.

If one is setting out to evaluate the effectiveness of one or more retrieval systems, one must clearly identify the unit or focal point of the analysis, the performance measure or measures to be used, and how these measures will be applied. Depending on one’s research objective, the unit of analysis may be, for example, search results, or user assessments of system ease of use. The two most common measures of retrieval system performance are recall and precision of search results, on the one hand, and user satisfaction with search results, on the other. Both measures were used in this study not only because they are the two most used measures of system performance and retrieval effectiveness, but also because research has shown that they are not always in agreement and that they may measure different things. Users of online library computer catalogs and other computer-based information retrieval systems often express satisfaction with their search results and the overall performance of the retrieval system even when the results, upon analysis, are shown to be poor. This has been called the phenomenon of “false positives” in user assessments of search success (Applegate, 1993:...
Associations between these different variables, actual search performance and users' satisfaction with search results, were also examined.

The study described here featured an experiment that was designed to compare users' search performance and assessments of ease of use or search success after use of either a Web-based hypertext OPAC, or its conventional, text-based counterpart. The study took place in a university setting. The research questions addressed in the study include the following:

1. Are Web OPACs more effective retrieval systems than their conventional text-based counterparts?
2. What factors are better predictors of search success, user interface factors, or search task level of difficulty?
3. What associations exist between search performance and factors such as perceived ease of use and satisfaction with search results?
4. Will WebOPACs be rated as superior in ease of use and search performance independently of search task level of difficulty?
5. What associations may be discovered between actual search performance and personal characteristics such as gender, educational level, major area of study, previous OPAC or Web experience?
6. Who will benefit most from these new network-accessible, hypertext, GUI Web OPACs?

2. Recent OPAC Research

Many accounts of OPAC research have been published in the last 15-20 years. A recent "Special Topic" issue of the Journal of the American Society for Information Science (JASIS) covered state-of-the-art OPAC research (Beaulieu and Borgman, 1996). Citations to many of the best research studies can be found in the articles published in this special issue of JASIS. Hildreth reviews much of this research in an essay on old and new design models for online catalogs. (Hildreth, 1995a). An analytical review of recent OPAC research is provided by Large and Beheshti (1997). The authors focus on the various methodologies employed in OPAC studies, and summarize research-based recommendations under three headings: database record enhancement, search capabilities, and interface design. For an exhaustive review of the information retrieval system evaluation literature see: Harter and Hert, Evaluation of information retrieval systems: approaches, issues, and methods. (1997).

Little has been published to date on the performance of Web-based OPACs or user satisfaction with these GUI, hyperlink-capable online catalogs. A search for research on Web-based OPACs turned up only a few publications. Apparently, no experimental studies involving Web OPACs have been conducted. Hildreth (1995b) looked at the new graphical user interfaces (GUIs) that are being applied to older, conventional, second generation OPACs. He warned that users may be too easily impressed with these systems, systems that deliver the same old level of poor results. In her insightful 1996 article, Borgman asks, "Why are online catalogs still so hard to use?" (Borgman, 1996). Perhaps we should be asking now, "Why do easy OPACs still produce such poor results?"

The presentation of bibliographic information in Web OPACs has been the focus of several recent studies. Cherry reports on her comparative study of bibliographic displays in 12 conventional OPACs and 10 Web OPACs. She developed an index of desirable display features to permit scoring of the two sets of OPACs. When assessed against this checklist of features, Web OPACs scored only slightly higher than pre-Web OPACs (60 percent to 58 percent). (Cherry, 1998:124). Ayres, Nielsen, and Ridley (1999) describe "BOPAC2," a research project funded by the British Library Research and Innovation Centre that was designed to test and evaluate a Web front end. This front end facilitates uniform access to a number of different library catalogs via the Z39.50 search and retrieval protocol. The BOPAC2 (the "B" is for Bradford University) research focuses on the management at the user interface of very large and complex retrieval sets. Research is still in progress and updates are provided at this Website: http://www.comp.brad.ac.uk/research/database/bopac2.html (accessed 3/2/00). Carlyle and Timmons (1999) have recently completed a comparative study of the composition of
default displays of bibliographic records in 100 Web OPACs. A report of this study can be found on the Web at: http://www.ischool.washington.edu/research/projects.cfm (accessed 3/4/00).

In "Web-based OPACs: Between Tradition and Innovation," Ortiz-Repiso and Moscoso (1999) report on their analytical study of Web OPACs. The goal of their research is to investigate OPACs available on the Web in order to ascertain how successful these new OPACs are at solving the problems associated with first and second generation OPACs. In these preliminary notes, the authors point out that in spite of significant enhancements to the user interface and the overall quality of interaction with the OPAC systems - most notably hypertext search and browse features, many previously documented problems experienced by users of first and second generation OPACs are still encountered in Web OPACs. They attribute this to the fact that the underlying structure of the Web OPACs remains unchanged (e.g., record format and content, indexing, and search algorithms).

3. Project Methodology

Overview

To address the research questions articulated in this study, and to test several research hypotheses, an experiment was designed to compare university students' use and perceptions of two online library catalogs (OPACs), a Web OPAC, and the conventional version of the same OPAC. A randomized, multi-factor block design was used. Four independent test groups were defined by the combination of the two experimental factors, OPAC interface style (Web-based OPAC and Text-based OPAC), and level of search task difficulty (easy and hard - See Table 1). The random assignment of volunteer subjects to these experimental blocks minimized variations between test subjects and testing conditions. Furthermore, this design permits either independent factor, interface style or search task level of difficulty, to be considered as the “treatment” factor. The operational online catalog at the University of Tulsa was used for this experiment. The two OPACs differed only in user interface style and interaction capabilities. One version of the OPAC employed the conventional text-based command and menu-driven interface. The other version, the Web OPAC, featured a “point and click” graphical user interface (GUI). Hypertext searching and browsing was supported in the Web version. Both OPACs contained the same catalog database and searchable indexes. Both versions of the OPAC were accessible via the Internet. The Web OPAC was available on the World Wide Web, and the text-based OPAC was available via Telnet access on the Internet.

<table>
<thead>
<tr>
<th>INTERFACE</th>
<th>SEARCH TASKS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Easy</td>
</tr>
<tr>
<td>Text/Menu</td>
<td>Group-1 n=16</td>
</tr>
<tr>
<td>GUI/Web</td>
<td>Group-2 n=16</td>
</tr>
</tbody>
</table>

Table 1. OPAC Experimental Design: Group Characteristics

Participants in the study were recruited from the general undergraduate population of the University of Oklahoma. As volunteers arrived for their scheduled one-hour appointments, they were assigned serially to one of the four test groups, and this continued until a total of 16 subjects completed the test in each of the four experiment groups. The session monitor, either a research assistant or the principle investigator, welcomed the student to the study and the online catalog workstation. The workstation was already logged on to the appropriate OPAC. To minimize distractions during the search sessions, the workstation was located in a quiet isolated area of the university library building. No training in the use of the test OPAC was
given to the participant. The monitor handed the student participant a clipboard containing the search task worksheet and the post-search questionnaire. After a brief introduction to the project, the monitor left the room but remained on call nearby to assist with any technical problems that might arise. In only one case was it necessary to assist the participant to log back on to the appropriate OPAC.

Data Collection Instruments

During the test session, each participant was required to perform a set of pre-selected search tasks, and to record the results of the searches on the search worksheet. With assistance from professional librarians and experienced database searchers, two sets of OPAC search tasks were designed, a set of easy searches, and a set of hard searches. Three searches were selected for the easy set, and four searches for the more difficult set. With two different OPACs and these two sets of search tasks, four experimental groups were defined by a unique combination of OPAC and search task level of difficulty. For purposes of simplification in the remainder of this report, each of the four test groups will be labeled in accord with its OPAC and task level. Group 1 will be “Text/Easy,” Group 2 – “Web/Easy,” Group 3 – “Text/Hard,” and Group 4 – “Web/Easy.”

The experiment took place over a period of four weeks. A few participants completed their searches in less than one hour. Those who had not finished searching after one hour were given a few more minutes to complete a search in progress. After the searches were completed, each participant was required to complete a post-search questionnaire. All 64 participants made some attempt to perform each of the searches in their designated search task set. All 64 participants completed the post-search questionnaire.

Each participant was instructed to report the results of a search by recording the found item’s exact call number on the search answer sheet. These call numbers were used later to score the participant’s actual search success for data analysis and group comparisons. The post-search questionnaire consisted of 17 questions. Questions 1 through 11 were designed to investigate users’ perceptions of system ease of use and usefulness, and users’ level of satisfaction with search results. Various dimensions of ease of use and searcher satisfaction were addressed by eleven statements. Participants were asked to respond to these statements on 4-point or 5-point Likert scales. Questions 1 through 7 were designed to create an ease of use measurement index. Six additional questions solicited data about the participant’s age, university status, major area of study, and previous experience with OPACs and the Web. The gender of each participant was recorded on the questionnaire by the researchers.

Independent and Dependent Variables

In summary, this experiment focused on the following variables: two independent variables, interface style (a WebOPAC and a TextOPAC) and search task level of difficulty, and four dependent variables, actual search performance, perceived ease of use, perceived system usefulness, and perceived search success (i.e., expressed satisfaction with search results).

Research Hypotheses

The following are among the research hypotheses tested in this study. As a convention in this report, the two catalogs will be referred to as the “WebOPAC” and the “TextOPAC.”

1. Search performance by users of the WebOPAC will be significantly superior to the search performance of users of the TextOPAC.
2. The search performance of users who perform easy search tasks will be superior to the search performance of users who perform difficult searches on both the TextOPAC and the WebOPAC.
3. Searchers actual search performance will be reflected in their level of satisfaction with search results.
4. More than actual search success and performance, perceived ease of use will influence searchers’ satisfaction with search results.
5. Users will judge the WebOPAC as superior in ease of use.
6. Satisfaction with search results will be greater among users of the WebOPAC.

4. Data Analysis and Findings

Personal Characteristics

The data collected on the personal characteristics of the 64 participants in the study are displayed by group in Tables 2a and 2b. For four of these variables, university status, number of OPACs used, catalog use frequency, and Web searching frequency, original response choices were consolidated for purposes of data analysis. For example, “University Status” was collapsed from the four categories, Freshman, Sophomore, Junior, and Senior, to lower division and upper division. The number of responses were combined, accordingly. Five response categories for “Catalog Use Frequency” were collapsed into just two, Low and High (High = “Once or Twice a Week” plus “Daily or Almost Daily”).

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>Mean Age</th>
<th>Univ Status</th>
<th>Major</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>M</td>
<td>L</td>
<td>U*</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>5</td>
<td>23.4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>8</td>
<td>20.8</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>8</td>
<td>22.3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>8</td>
<td>21.4</td>
<td>4</td>
</tr>
<tr>
<td>Totals</td>
<td>35</td>
<td>29</td>
<td>21.97</td>
<td>17</td>
</tr>
<tr>
<td>%</td>
<td>54.7</td>
<td>45.3</td>
<td>26.6</td>
<td>18.8</td>
</tr>
</tbody>
</table>

Table 2a. User Profile By Group: Personal Characteristics

(*Note: Univ. Status ‘L’=Freshmen/Sophomore; ‘U’=Junior/Senior Major ‘H’=Humanities, ‘N’=Natural Sciences, ‘S’=Social Sciences)

<table>
<thead>
<tr>
<th>Group</th>
<th>OPACs Used</th>
<th>Catalog Use Freq</th>
<th>Web Use Freq</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Few</td>
<td>Many</td>
<td>Low</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Totals</td>
<td>15</td>
<td>49</td>
<td>46</td>
</tr>
<tr>
<td>%</td>
<td>23.4</td>
<td>76.6</td>
<td>71.9</td>
</tr>
</tbody>
</table>

Table 2b. User Profile By Group: Online Search Experience

Several pertinent facts are revealed in this profile of participants in the study. Most participants were upper division undergraduate students (73.4 %), and most were majoring in one of the natural sciences or mathematics (57.8 %). With regard to search experience, most had used at least 2-3 different library computer catalogs, but most of these reported low frequency of library catalog use (71.9%). On the other hand, all but three of the participants reported high use of a Web search engine or search service to look for information on the Internet.

Search Performance Scoring

Participants were required to record the call numbers of retrieved items on the search task answer sheet. The researchers conducted extensive and repeated searches of the Tulsa University database to discover and retrieve all the relevant or possibly relevant records for each test search. If an item recorded on a participant’s answer sheet was not among these records (a rare occurrence), the full bibliographic record was retrieved to assess the relevance or non-relevance of the item.

Each of these found items was judged by the researchers on a 3-level scale of relevant, possibly relevant, and not relevant. Relevant items received 2 points, possibly relevant items received 1 point, and non-relevant items received no points. Each searcher’s points were totaled and then transformed to an equivalent number on a 100 point scale. This trans-
formation adjusted for the different amounts of total points achievable on the hard search task sheet (20), as opposed to the easy search task sheet (16). This adjustment made it possible to meaningfully compare the mean search scores of all four groups. The mean search performance scores of each of the four groups are presented as the first value in each cell of Table 3.

<table>
<thead>
<tr>
<th>SEARCH TASKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy</td>
</tr>
<tr>
<td>Text/Menu</td>
</tr>
<tr>
<td>81.641</td>
</tr>
<tr>
<td>13.750*</td>
</tr>
<tr>
<td>22.250**</td>
</tr>
<tr>
<td>GUI/Web</td>
</tr>
<tr>
<td>83.984</td>
</tr>
<tr>
<td>15.313*</td>
</tr>
<tr>
<td>23.250**</td>
</tr>
</tbody>
</table>

Table 3: Group Mean Values: Search Performance Scores; Search Results Satisfaction*; Ease of Use Rating**

Post-Search Questionnaire Responses

Several of the questions on the post-search questionnaire were designed to gather data on users’ satisfaction with their search results, and assessments of system ease of use. In question 10, “How satisfied were you with the results of your searches?” participants responded on a five-point Likert scale, ranging from “Very satisfied” to Very unsatisfied.” If one assumes a roughly equal distance between each of the values on this five-point scale, the mean values of the four groups shown in Table 3 indicate little or no difference in searchers’ level of satisfaction with their search results. (For purposes of analysis, this ordinal scale was transformed into an ordinal scale as follows: Very satisfied = 20 points, Satisfied = 15 points, and so on.) A closer look at the responses reveals that 37.6 percent of the Group 3 (Text/Hard) participants expressed no more than a “3” level of satisfaction with their search results.

Questions 1-7 were designed to measure several specific aspects of OPAC system ease of use. Question 8 was designed to measure overall ease of use. Each of the questions 1-7 required participants to indicate their level of agreement on a 4-point scale with a statement that described some aspect of interaction with the OPAC. A 4-point scale was used to eliminate fence-sitters. Question 8 asked subjects to assess the “overall ease of use” of the OPAC on a 5-point scale, from “Very difficult” to “Very easy.”

It has become common practice in social sciences research to construct multi-factor measurement indices by combining several questions that address related aspects of a single variable (e.g., ease of use). (Sirkin, 1995, 68). While each question may use an ordinal response scale, taken together, the responses to these questions are transformed into a single composite interval level of measurement index. As Schutt explains, “When several questions are used to measure one concept, the responses may be combined by taking the sum or average of responses. A composite measure based on this type of sum or average is termed an index or scale.” (Schutt, 1999, 75). And Schutt continues: “In addition, the index can be considered a more complete measure of the concept than can any one of the component questions.”

Questions 1-7 were combined into a single interval level index to provide a better, more accurate measure of a participant’s assessment of the OPAC’s ease of use. Each “Strongly agree” response was assigned 4 points, each “Agree” response 3 points, each “Disagree” response 2 points and so on. A score total of 28 would reflect the strongest agreement on all seven questions. The results of this indexed ease of use measure are provided as the last value in each cell of Table 3. The responses to question 8, “overall ease of use” appear to support the validity of this index.
Independent Group Comparisons

To compare the mean search performance scores and rating assessments of the four groups, and to test the research hypotheses, both parametric (independent group, two-sample t-tests) and non-parametric (Mann-Whitney tests) statistical procedures were employed in the analysis of the data presented here. When the data are not normally distributed, the Mann-Whitney tests produce more valid results.

Comparisons of each pair of the test groups were conducted to discover any statistically significant differences in the values of the three dependent variables. Such differences would support the rejection of one or more null hypotheses (no difference) regarding these variables, and provide support for the research hypothesis.

Search Performance Scores

Each group's search performance scores were compared to every other group's scores. Of the six possible pairings, significant differences were found between Group 1 (81.641) and Group 3 (70.625, p=0.024), as well as between Group 2 (83.984) and Group 3 (70.625, p=0.006). (In all these comparisons, the level of significance chosen is 0.05.) The Mann-Whitney tests support these findings, and also suggest a significant difference exists between the performance scores of Group 2 (83.984) and Group 4 (75.625, p=0.037). WebOPAC searchers significantly outperformed the TextOPAC searchers, but easy task searchers also outperformed hard task searchers on both OPACs.

Satisfaction with Search Results

After a glance at Table 3, one might conclude that there was no appreciable difference among searchers in the four groups when it came to expressing satisfaction with their search results. The results of the t-tests and Mann-Whitney tests confirm this conclusion. No significant differences were identified. This suggests a disjunction between actual search results and subjective assessments of the quality of those results. Those participants with the lowest search performance scores express equal or nearly-equal levels of satisfaction with search results as those with much higher search scores.

Ease of Use Assessments

Questions 1-7 on the post-search questionnaire comprised the ease of use index which made possible an interval level of measurement. The mean rating scores of the four test groups are displayed in Table 3. The most significant difference in user assessments of OPAC ease of use exists between Group 2 (23.250) and Group 3 (20.250, p=0.010). A significant difference in ease of use assessments was also discovered between Group 3 (20.250) and Group 4 (22.313, p=0.040). Mann-Whitney tests strongly supported the t-test results for Groups 2 and 3 (p=0.006), but provided only weak support for the Groups 3 and 4 finding (p=0.068). Both WebOPACs received higher ease of use ratings than their text counterparts.

Type of OPAC and Task Level Group Comparisons

The four original test groups can be realigned to create four new groups for purposes of analysis. The four new groups include, respectively, all participants who used the TextOPACs, all those who used the WebOPACs, all those who performed the easy search tasks, and all those who performed the hard search tasks. The mean values for these groups, so considered, are displayed in Table 4.

Immediately it can be seen that the highest level of search performance was achieved on the easy search tasks, while the lowest level of performance was registered by those required to carry out the hard searches. WebOPACs received higher ease of use ratings than TextOPACs. The superiority in search performance achieved by the easy task searchers over
the hard task searchers is significant (p=0.005). WebOPAC searchers' performance scores were higher than the performance scores of TextOPAC searchers, but this difference was not found to be significant.

<table>
<thead>
<tr>
<th></th>
<th>Mean Search Scores</th>
<th>Mean Results Satisfaction</th>
<th>Mean Ease of Use Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>TextOPACs</td>
<td>76.1328</td>
<td>13.7500</td>
<td>21.2500</td>
</tr>
<tr>
<td>WebOPACs</td>
<td>79.8047</td>
<td>14.5313</td>
<td>22.7813</td>
</tr>
<tr>
<td>Easy Tasks</td>
<td>82.8125</td>
<td>14.5313</td>
<td>22.7500</td>
</tr>
<tr>
<td>Hard Tasks</td>
<td>73.125</td>
<td>13.7500</td>
<td>21.2813</td>
</tr>
</tbody>
</table>

Table 4. Group Comparisons by OPAC and Task Level

Assessments of ease of use appear to be affected by search task level of difficulty. Easy task searchers rated their OPACs higher in assessments of ease of use. The higher ease of use rating given by the easy task searchers was statistically significant (t-tests: p=0.05; Mann-Whitney: p=0.033). Additional t-test analysis demonstrated that the ease of use ratings provided by the TextOPAC users and the WebOPAC users are significantly different (p=0.041). The WebOPAC was rated superior in ease of use. Regarding satisfaction with search results, no significant group differences were identified.

Associations Between Dependent Variables

Pearson's correlations coefficient (r), a parametric procedure, and Spearman's Rank correlation coefficient, a non-parametric procedure, were used to examine the relationships between the dependent variables, interpreted as quantitative variables. No significant associations were found between search performance scores and either of the dependent variables (ease of use and satisfaction with search results). A strong positive correlation was identified between the satisfaction with search results ratings and the ease of use assessments (r=0.5992, p<0.001). This analysis indicates that actual search performance is not a predictor of a user's satisfaction with search results. It does indicate that users' perceptions of ease of use exert a stronger influence on satisfaction with results than actual search performance.

Tests for Independence

Tests for independence (Chi Square) were conducted to examine whether participants grouped by type of OPAC used or by level of search task difficulty responded differently to questions 1-9 (interpreted as category responses) on the post-search questionnaire. This analysis revealed a significant association between participants' responses to question 5, “My parents could use this library computer catalog to search effectively with little or no training”, and the type of OPAC used (p=0.005). A significantly greater number of WebOPAC users agreed with this statement than did TextOPAC users. WebOPAC users seem to be saying, "WebOPACs are so easy to use, even my parents could use them effectively with little or no training!"

Associations Between Personal Characteristics and the Dependent Variables

Chi Square tests for independence were conducted to identify any significant associations between the personal characteristics and either of the dependent variables, satisfaction with search results and ease of use. No associations were found between any of these variables and gender, subject major, number of OPACs used, or catalog use frequency. An association was discovered between university status and ease of use (p=0.010). Upper division students rated their OPACs ease of use much higher than lower division students rated theirs. An association was also discovered between frequency of Web searching and satisfaction with search results (p=0.033). Frequent Web searchers expressed higher levels of satisfaction with search results than infrequent Web searchers.
5. Summary and Conclusion

The discussion of the findings in this study focused on three areas: interactions between the independent variables (interface style and task level of difficulty) and three dependent variables (search performance, satisfaction with search results, and ease of use); associations between dependent variables; and associations between personal characteristics and dependent variables.

Upon reflection on these findings, certain key observations come to mind. Both interface style and search task level of difficulty exert an influence on search performance and ease of use factors. Performing the same search tasks, WebOPAC users scored higher than Text-OPAC users, but the difference was not statistically significant at the .05 significance level. Using the same OPAC, easy task searchers outperformed hard task searchers by a wide margin. The difference was found to be significant. The magnitude of this difference suggests that task level of difficulty exerted a stronger influence on performance scores than interface style.

WebOPACs were rated higher than TextOPACs in ease of use. The difference was found to be significant. Easy task searchers rated their system's ease of use higher than hard task searchers, but this difference is not significant. No significant differences were found between any of the groups in participants' satisfaction with search results.

Few associations were found between personal characteristics and the dependent variables. Upper division students rated both OPACs easier to use than lower division students. A significant association was discovered between frequency of Web searching and satisfaction with results. Frequent Web searchers expressed much higher levels of satisfaction with search results. They also were more likely to use WebOPACs again than infrequent Web searchers. Web use seems to lead to more Web use.

The results of the correlation analysis are the most telling. Once again, the evidence indicates that actual search performance is not a predictor or determinant of a searcher's satisfaction with search results. There seems to be little interaction between these two variables. On the other hand, perceptions of ease of use do influence users' satisfaction with search results.

In summary, when considering search performance, search task level of difficulty seems to be a major determinant, but OPAC interface style may affect search performance as well. When considering perceptions of ease of use, interface style appears to be the primary determinant. Furthermore, as earlier studies have shown, there is little or no association between actual search performance by users and their expressed satisfaction with search results. In short, WebOPACs are easier to use, and this may be a supportive factor in search success. However, the down side is this: WebOPACs may contribute to “false positives” in users' assessments of search results, and this may explain, at least partially, why users are often satisfied with poor search results.

Participants selected themselves to participate in this study. For this reason, caution must be exercised in extending these findings to any larger population of OPAC users. The external validity of the study has not been established. Participants were randomly assigned to the test groups, but this aggregate of volunteer participants constitutes a convenience sample, not a true probability sample.

Assigning numerical scores to search results is always a risky business. The scoring of results in this study was based on “relevance” judgments made by the searchers and the researchers. These judgments were made after viewing only standard catalog records, well-known for their paucity of “content” information that is helpful in such judgments. The scoring of search results was done as consistently as possible, and the total score of each searcher was derived from the items found and recorded. Relevant items they did not retrieve did not factor into these computations. Nonetheless, this searching took place in an “artificial” setting where a searcher's actual information needs and search objectives were not factored
into the search attempts or the relevance assessments. Judgments of relevance made in such settings are understandably open to question.

Serious questions remain about the validity and reliability of satisfaction with search results as a measure of system performance. This research has shown that satisfaction with search results is often influenced by non-performance factors like perceived ease of use. More testing and evaluation of satisfaction measurement indexes is needed. It will be a continuing challenge for researchers to devise refined measures of user satisfaction with search results and system performance.

6. References