Purdue University
Purdue e-Pubs

Proceedings of the IATUL Conferences   1996 IATUL Proceedings

Testing FirstSearch and UMI Databases in Technological University Libraries

Charles Townley
New Mexico State University

Towards a Performance Based Model
for Selecting Electronic Information

Townley, Charles T.
New Mexico State University, USA

Introduction

In the 1960's, librarians began using electronic databases as a part of library services. In the beginning, electronic databases were usually searched by professional staff as an exceptional reference service. In the 1980's, libraries began using CD-ROM and locally mounted tape versions of electronic databases. In recent years, Internet access and consortial buying have diversified availability of electronic information. Indeed, a growing number of libraries now have multiple ways to access electronic information.

As electronic information and access have grown, selection issues have become more complex. When alternatives were highly limited, selection was based primarily on access and cost. As CD-ROM and tape-loaded electronic information became available, local area networks and interfaces became important issues in selection. Now with multiple sources, human, demographic, and technological factors are all important. Factors like training opportunities, Z39.50 standards, password protection, links-to-holdings, and full-text availability are part of the selection decision. Because of the dynamic nature of electronic information, traditional selection criteria are not effective and new criteria must be developed.

Performance based selection criteria offer promise. Performance based criteria are drawn from factors which affect usage. Also, performance based criteria are dynamic and sensitive to rapid change. To the extent that key factors can be identified and measured, it is possible to determine which database and which access technology are most effective for a local library. This results oriented approach to selection permits constant fine tuning. And, in this time of limited resources, it helps keep the focus on the customer.

Library literature is full of reviews and comparisons of specific electronic resources and access technologies. But, most of this material provides no empirical information on selecting among alternatives. In the U.S., only Richard Meyer at Trinity University has proposed an outcomes based model for selection of electronic information. [1] This model recognizes the importance of the cost to the user and the library. It also identifies factors such as interface friendliness, links-to-holdings, and full-text availability as critical factors in improving the effectiveness of particular databases.

The impetus for this particular study came from the need to make a management decision at the NMSU library. In 1994, the library had just purchased tape-loaded versions of two general bibliographic databases. It had also built a CD-ROM LAN for use of CD-ROM databases in the library and on campus. It was our good fortune to have the opportunity to participate in a group purchase of FirstSearch along with other colleges and universities in New Mexico and Texas. After careful consideration, the University Library decided not to participate during the first year. We also stated that we would use the year to collect data comparing the performance of different access

technologies and databases. The result is following study on performance based selection of electronic information.

Study Design

The study is composed of two parts. The first part, described in this article, identifies the extent to which there are statistically significant differences in use of electronic databases among selected participating institutions. The first part of the study is intended to determine the extent to which performance of particular databases differs among institutions.

The second part of the study is intended to establish causality. It will use regression and other inferential statistical analysis to determine the percent of variance explained by technical, demographic, and human factors. It is expected that this study will be completed by September, 1996.

The study is based on data from six member libraries of the Llano Escadado Information Access Network (LEIAN) and the New Mexico Consortium of Academic Libraries (NMCAL) in the states of New Mexico and Texas. Five of the libraries in this study participate in FirstSearch through a joint contract administered by Texas Tech University and operated through the Amigos Bibliographic Council. New Mexico State University provides either tape mounted or cd-rom access to these same databases.

The institutions selected to participate are purposely diverse. They include:

- Abilene Christian University, a private liberal arts college.
- Eastern New Mexico University, a regional university.
- New Mexico Highlands University, a regional university.
- New Mexico State University, a research land grant university.
- Texas Tech University, a research university.
- University of New Mexico, a research university.

Data was collected from three sources. Frequency counts were collected for use of Periodical Abstracts; Article First; Newspaper Abstracts; ABI Inform; ERIC; GPO Monthly Catalog; MLA Bibliography; and PsychInfo. Data was collected for the period of September, 1995 through April, 1996, the academic year. Counts were made using FirstSearch's definition of a search-entry of a line of search terms. In situations where this was not available, the frequency of database openings was multiplied by three. Also, Periodical Abastracts and ABI Inform counts for Texas Tech are reported with a combined count and divided by 45 and 55 percent respectively. These changes conform to a use-based conversion formula developed by Tex-Share to achieve equivalent counts. [2]

Second, enrollment statistics by major and credit hour were requested and provided by the institutions. Total enrollment was taken from Petersen's 1996 Guide. [3] Demographic information on majors and credit hours were provided by the institutions.

Finally, telephone interviews were held with the individual responsible for operation of electronic information systems at each institution. Data was collected regarding: interfaces; menus; campus networks; off-campus access; pass word protection; length of use; other sources for specific databases; printing capabilities; down-loading capabilities; instructional support; full-text availability; interlibrary loan access; and links to holdings. These three sources of data are used throughout the study.

The hypothesis for the study is that there are systematic differences in use as measured by:

- use per enrolled student
- use by majors, both undergraduate and graduate
- use per credit hour

Chi-square analysis is used to test null hypothesis of equal distribution. Statistically significant chi-squares values indicate systematic differences in usage patterns within a database.

General Sources

Four general indexes were included in this study: Article First; Periodical Abstracts, U.S. Government Publications Office Monthly Catalog, and Newspaper Abstracts. The indexes were selected for their breadth of subject scope, broad coverage of a particular format (i.e. periodicals, U.S. government publications, and major newspapers). They were also selected for their intensive use. Within the LEIAN Consortium, Article First accounted for 19.7 percent of use and the four databases combined accounted for more than 27 percent of use.

Chi-square analysis was used to determine where statistically significant differences existed in usage patterns of the indexes across institutions. [4] The chi-square test is a non-parametric test of distributions. The test determines the probability that a sample distribution is the result of drawing from a population with a theoretical (in this case, equal) distribution. If the results are statistically significant, the distribution is not considered even and the null hypothesis is rejected. For this study, rejecting the null hypothesis supports the hypothesis that there are systematic differences in use of databases at the different institutions.

Article First and Periodical Abstracts counts were combined by institution. This was done to achieve an equivalent count of general periodical index use among the institutions during the period covered by this study. The results indicate a chi-square of 160,573 with five degrees of freedom, indicating a significance level in excess of .001 (Table 1). Statistically, there are significant differences among the usage patterns of the different institutions. While the characteristics of use are not being studied in the first part of the study, it is interesting to note that low use by Abilene Christian University, Eastern New Mexico University, New Mexico Highlands University, and the University of New Mexico may be closely related to the fact these databases form the second line of general periodicals inquiry at these institutions. Other databases provide primary access to general periodicals. Abilene Christian University uses locally mounted Wilson indexes. Eastern New Mexico University and New Mexico Highlands University use Info Track as their first point of inquiry for general

periodical use. The University of New Mexico offers a locally mounted version of Expanded Academic Index as well as a gateway to Uncover.

	Total	1996	Use/
Institution	Use	Enrollment	Student
Abilene Christian College	2268	4207	.54
Eastern NM	5222	3853	1.36
NMHU	2998	2797	1.07
NMSU	102993	15645	6.58
Texas Tech	164170	24083	6.82
UNM	14350	24344	.59
Total	292001	74929	3.90

TABLE 1. PERIODICAL ABSTRACTS & ARTICLE FIRST USE BY STUDENT
X2 = 160,573 5dF s.@.001

Perhaps more interesting is the small difference in per capita use between Periodical Abstracts locally mounted on the Opac at New Mexico State University and combined Periodical Abstracts and Article First usage at Texas Tech. This raises the question of the cost effectiveness of locally mounted databases.

The U.S.G.P.O. Monthly Catalog index indexes government publications. In this case all the libraries analyzed, except NMSU, use FirstSearch as their primary means of accessing the U.S.G.P.O. Monthly Catalog. NMSU uses a compact disk version on a CD-ROM LAN available in the library and on campus. Searches for NMSU are estimated by multiplying the number of databases opened by three, a conservative figure for estimating the number of searches each time a database is opened.

The chi-square statistic for the U.S.G.P.O. Monthly Catalog is 12,894 with four degrees of freedom, indicating a significance level in excess of .001. FirstSearch access to the U.S.G.P.O. Monthly Catalog is the second means of access to government documents at Abilene Christian University, Eastern New Mexico University, Texas Tech, and the University of New Mexico. Abilene Christian University uses a Marchive CD-ROM for primary access. Texas Tech and the University of New Mexico all have Marchive cataloging loaded on their Opac. This would explain the relatively high use at New Mexico State, where Marchive is only partially loaded into the NMSU catalog.

	Total	1996	Use/
Institution	Use	Enrollment	Student

Abilene Christian College	1217	4207	.29
Eastern NM	746	3853	.19
NMSU	9216	15645	.59
Texas Tech	4228	24083	.18
UNM	1079	24344	.04
Total	16486	72132	.23

TABLE 2. U.S.G.P.O. MONTHLY CATALOG USE BY STUDENT
$X2 = 12,894$ 4dF s.@.001

Other, more subtle differences, may also be present. For example, usage compared with the general periodical usage is much lower. Is this a format issue, a links-to-holdings issue or what? The second phase of this project will address these factors.

Newspaper Abstracts usage generates a chi-square of 67,619 with four degrees of freedom. This is statistically significant at the .001 level (Table 3). Both Eastern New Mexico and New Mexico Highlands University have access to similar information on their InfoTrack cd-rom systems. Abilene Christian University does not use the Newspaper Abstracts service at all, relying on full text from Lexus Nexus.

	Total	1996	Use/
Institution	Use	Enrollment	Student
Eastern NM	310	3853	.08
NMHU	297	2797	.11
NMSU	32622	15645	2.09
Texas Tech	9119	24083	.38
UNM	2255	24344	.09
Total	44603	70722	.63

TABLE 3. NEWSPAPER ABSTRACTS USE BY STUDENT
$X2 = 67,619$ 4dF s.@.001

The comparatively high use at New Mexico State University may be due to the fact that it is on the NMSU Opac and shares the same interface. In later analysis, this may also help determine the cost effectiveness of locally mounted databases.

Disciplinary Sources

Four disciplinary databases were addressed: ABI Inform; ERIC; Modern Languages Association (MLA) Bibliography; and PsychInfo. Data were collected on undergraduate and graduate majors by full-time equivalent units and on credit hour production. The use data are cleaner in the disciplinary databases, being based on FirstSearch statistics only at Abilene Christian University, Eastern New Mexico University, Texas Tech University, and the University of New Mexico. NMSU data

was database opening data from the local area network cd-rom server only multiplied by three to approximate the actual number of searches. With the exception of ERIC, access to these databases is limited to only one access technology.

Chi-square analysis indicates that usage patterns for ABI Inform were significantly different, both for enrolled student FTE and credit hours (Table 4). Since an alternative source was not readily available, other factors must be significant in explaining this difference. Interview data suggests that perhaps placement on the menu or emphasis during bibliographic instruction may explain some of the difference.

Institution	Total Use	Total FTE	Credit/Hr Production	Use/Total FTE	Use Credit Hr
Eastern NM	1751	358	5303	4.89	.33
NMSU	21792	1894	22283	11.51	.98
Texas Tech	115500	2236	32491	51.65	3.55
UNM	4095	1300	15749	3.15	.26
Total	143138	5788	75826	24.73	1.89

TABLE 4. ABI INFORM USE BY MAJOR AND CREDIT HOUR
Major X2 = 109,382 3dF s.@.001
Credit Hour X2 = 58,990 3dF s.@.001

Another interesting observation is that uses per credit hour in a professional program like business are high compared with academic disciplines like psychology or English, languages and linguistics, anthropology, and archeology found in PsychInfo and MLA databases. Is this reasonable? And if so, what does it mean?

ERIC data was statistically significant (Table 5). The ubiquitous nature of ERIC in the U.S. means that there are numerous other means of accessing it at each institution. Abilene Christian University supports some out-of-date CD-ROMs for retrospective searching. Eastern New Mexico maintains a current cd-rom subscription. The College of Education at New Mexico State University maintains a second ERIC CD-ROM subscription in its Educational Materials Laboratory. Texas Tech has a stand alone ERIC CD-ROM product. And, the University of New Mexico has access to ERIC through CARL. Locally imposed printing limitations at Eastern New Mexico University and heavy use of CARL at the University of New Mexico may suggest interface limitations with FirstSearch.

Institution	Total Use	Total FTE	Credit/Hr Production	Use/Total FTE	Use Credit Hr
Abilene Christian College	5765	563	2579	10.24	2.24

Eastern NM	3725	554	8099	6.72	.46
NMSU	34779	1735	14706	20.05	2.36
Texas Tech	21045	939	12887	22.41	1.63
UNM	5893	1876	32976	3.14	.18
Total	71207	5667	71247	12.57	1.00

TABLE 5. ERIC USE BY MAJOR AND CREDIT HOUR
Major X2 = 29,944 4dF s.@.001
Credit Hour X2 = 103,813 4dF s.@.001

Modern Language Association's (MLA) Bibliography usage is significantly different, both for majors and for credit hour production. Results are shown in Table 6. Online access to other sources, in this case Humanities Index, does not seem to negatively affect comparative use at Abilene Christian University. Eastern New Mexico University indicates that local printing limitations may discourage use. The higher use by major at New Mexico State University may relate to the more effective interface available on compact disk compared with FirstSearch. On the other hand, the lower credit hour production number at NMSU may have to do with something outside the scope of this studies, such as the number of required introductory English courses.

Institution	Total Use	Total FTE	Credit/Hr Production	Use/Total FTE	Use Credit Hr
Abilene Christian College	2883	131	4653	22.01	.62
Eastern NM	1205	128	1832	9.41	.66
NMSU	4785	334	21056	14.33	.23
Texas Tech	24435	2258	33459	10.82	.73
Total	33308	2851	61000	11.68	.55

TABLE 6. MLA BIBLIOGRAPHY USE BY MAJOR AND CREDIT HOUR
Major X2 = 1,577 3dF s.@.001
Credit Hour X2 = 6,089 3dF s.@.001

Use of the PsychInfo database is significantly different (Table 7). The comparatively high use at New Mexico State University may be result of the comparatively large number of graduate students or a psychology collection which is perceived as being weak, an issue outside the scope of this study.

Institution	Total Use	Total FTE	Credit/Hr Production	Use/Total FTE	Use Credit Hr
Abilene Christian	5851	213	2762	27.47	2.12

College					
Eastern NM	2436	142	2075	17.15	1.17
NMSU	24951	476	9340	52.42	2.67
Texas Tech	30449	757	11035	40.22	2.76
Total	63687	1588	25212	40.11	2.53

TABLE 7. PSYCH INFO USE BY MAJOR AND BY CREDIT HOUR
Major X2 = 4,475 3dF s.@.001
Credit Hour X2 = 1,993 s.@.001

Conclusions

The first stage of this study clearly establishes that use of electronic databases is significantly different among institutions of higher education in the southwestern United States. From a statistical point of view, the conclusion is overwhelming, given the high frequencies which make the chi-square test very robust. On the other hand, there is no systematic differentiation by type of institution. Some databases are used less frequently by research university than by regional universities. The private institution does not seem to make markedly higher use than the public schools. A larger sample of institutions might, however, establish some patterns by type of institution.

For the purpose of this study, the usage data clearly indicate that electronic information is used in very different ways at different institutions. Newspapers Abstracts, for example, is used much more heavily at New Mexico State than it is at the other institutions. And, the entire FirstSearch family appears to fulfill a different and secondary role at the University of New Mexico than it does at other institutions.

On the other hand, the similarity of usage patterns between Texas Tech and New Mexico State University, despite totally different access technologies, suggests that many differences in use may be the result of human or demographic factors.

Further analysis in the second part of this study will identify human, demographic, and technical issues that are affecting use within this study. We will be looking at human issues such as bibliographic instruction availability, length of use, and interface friendliness. We will be working with technical issues such as menu access, links to holdings, interlibrary loan request access, and printing. We will be dealing with demographic issues such as the undergraduate to graduate student ratio and length of use. From this analysis, it is hoped that we will be able to explain a significant amount of variance in usage patterns. This data can then be used to predict the effectiveness of different electronic information alternatives.

This performance based approach will not create "the" specific formula for evaluating electronic information in any venue. Nor will it establish importance of any particular factor in any particular institution. But it will provide guidance in how to weigh factors in making local selection decisions for electronic information. This will be quite useful as libraries continue to move into an environment of multiple databases and multiple technologies for accessing each database. It will provide the guidance for

meeting local information needs with electronic opportunities in the most cost effective manner.

References

1. MEYER, Richard W. **Selecting electronic alternatives. Information technologies and libaraies**. 11(2) 1993: pp.173-180.

2. Telephone conversation with Virginia Andrews, Texas Tech University, May, 1996.

3. **Perterson's register of higher education**. Princeton, Peterson's, 1996.

4. KOOSIS, Donald J. **Statistics**. New York, Wiley, 1972: pp. 215-234.