11 Failure Analysis of Subject Searching Capabilities

11.1 Introduction

Chapter 11 features a failure analysis of subject searching capabilities in the Blue and Pinstripe Systems. It focuses primarily on both the success and failure of these capabilities.

Table 11.1 lists Blue-system subject searches and the number of matches these capabilities made for user queries. Table 11.1 does not include matches that the Blue System made as a result of a user action to expand searches. It is divided into two portions: (1) Blue-system searches that made matches, and (2) reasons for the failure of the Blue System to effect matches.

Table 11.1. Matches in Blue-system Searches

Search approach	UM-D Blue	Earlham Blue
Searches making matches		
Exact	149	23
Alphabetical	64	12
Keyword-in-main-heading	6	0
Keyword-in-subdivided-heading	2	7
Title-keyword	52	15
Keyword in subject heading fields	2	2
Keyword-in-record (subjects)	6	5
Keyword-in-record (personal names)	0	4
Alphabetical (personal names)	0	19
Subtotal	281 (79.6%)	87 (83.7%)

No matches		
Vocabulary problems	16	10
Database failure	49	5
Search approach failures	4	2
Spelling problems	3	0
Subtotal	72 (20.4%)	17 (16.3%)
Total	353	104

Subject searching capabilities made matches for 79.6% and 83.7% of queries entered by Blue-system users at UM-D and Earlham, respectively. The exact, alphabetical, and title-keyword searches frequently made matches of user queries. Also, the Blue system's alphabetical approach frequently made matches of the personal-name queries that Earlham users entered.

The Blue System's subject searching capabilities failed to make matches for 20.4% and 16.3% of queries at UM-D and Earlham, respectively. Under the "no matches" portion of Table 11.1 are listed four reasons for search failure (i.e., vocabulary problems, database failure, search approach failure, and spelling problems). The discussion in Chapter 10 gave these same reasons for the failure of many more searches than are listed in Table 11.1. Chapter 10 failures were attributed to the failure of searches based on user relevance assessments (i.e., failed searches retrieved no titles that users rated "very useful"). Chapter 11 failures were attributed to the system's inability to produce retrievals in response to user queries.

Table 11.2 lists subject searching approaches in the Pinstripe System and the number of queries submitted to these approaches. It submitted a lot of queries for subjects generally to the keyword-in-record search.

Table 11.2. Queries Submitted to the Pinstripe System

Search approach	UM-D Pinstripe	Earlham Pinstripe
Searches making matches		
Alphabetical	53	11
Keyword-in-subdivided-heading	124	23
Keyword-in-record (subjects)	179	43
Keyword-in-record (personal names)	n/a	21
Total	356 (100.0%)	98 (100.0%)

11.2 Controlled Vocabulary Searching Based on the Exact Approach

11.2.1 Review of the Exact Approach

In the Blue System, controlled vocabulary searching focused on the exact approach. If the user-entered term matched an exact or normalized form of a controlled vocabulary term, the Blue System responded with the exact-approach main menu. The menu gave users several options for continuing their search: displaying titles assigned the matched term, displaying subdivisions connected with the matched term, browsing notes, browsing broader or narrower terms, or expanding their results with the results of a different subject searching approach.

The exact approach formed a large part of the Blue System's alphabetical approach and keyword-in-main-heading search. If the user-entered term matched a longer subject heading or *see* reference, the system responded with the alphabetical approach. This approach featured an alphabetical list of headings and references in the alphabetical neighborhood of the matched term. The system responded to users who chose a listed *see* reference with a message telling users that the system was retrieving titles assigned the particular subject heading connected with the *see* reference. It then constructed an exact-approach main menu for the particular subject heading. The system responded to users who chose a listed subject heading with the exact-approach main menu. Options in the exact-approach main menus were enabled or disabled depending on their availability for the selected subject heading.

If words in the user-entered term matched words in the main-heading subfield of subject headings, the Blue System listed matched main headings in alphabetical order. The system responded to users who chose a listed main heading with the exact-approach

main menu. Options in the exact-approach main menus were enabled or disabled depending on their availability for the selected main heading.

The alphabetical and keyword-in-main-heading approaches were little more than selection mechanisms for users whose entered terms did not match subject headings or *see* references in their entirety. When users made a selection, exact-approach options enabled them to explore terminology and display titles.

The alphabetical approach was also implemented in the Pinstripe System. A random selection algorithm governed the Pinstripe System's selection of the alphabetical approach. The Pinstripe System responded with the alphabetical approach regardless of the particular elements in user queries.

11.2.2 Precision in Controlled Vocabulary Searches

Figure 11.1 shows precision scores for the three different types of controlled vocabulary searches that were designed around the exact approach. (Chapter 11 figures showing precision scores include conflicting relevance assessments and relevance assessments for repeated displays of the same titles because these figures isolate retrievals from particular subject searching approaches.)

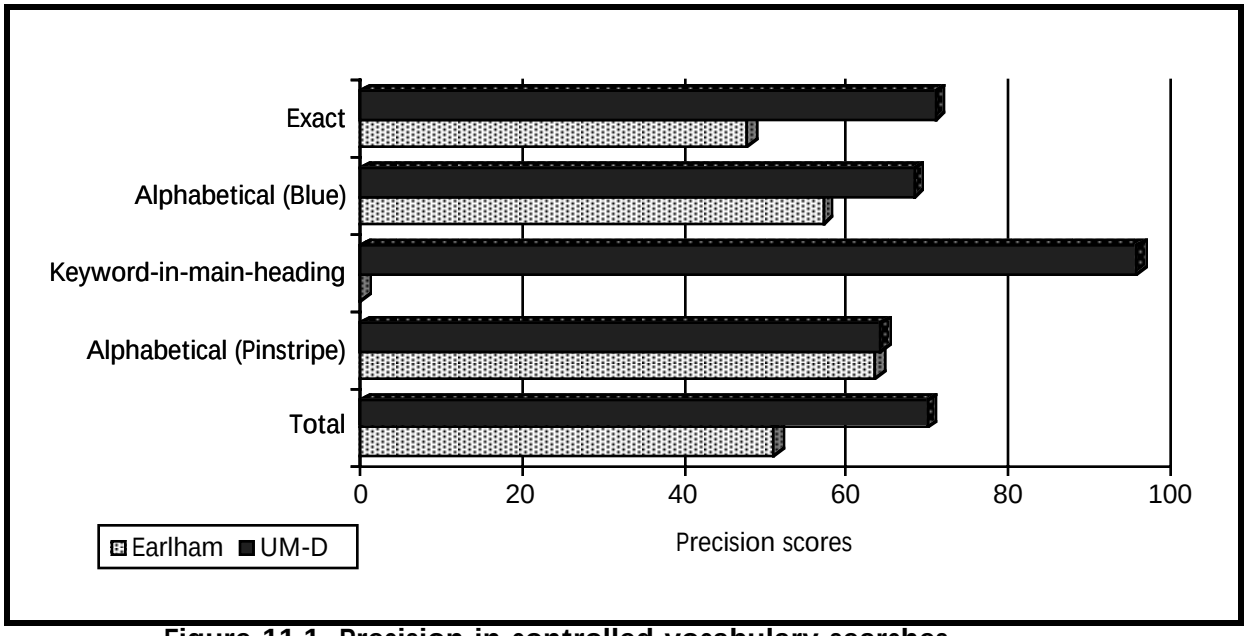

Figure 11.1. Precision in controlled vocabulary searches

Precision in controlled vocabulary searches varied considerably. They ranged from a low of 47.8% for exact searches in Earlham's Blue System to a high of 95.8% for keyword-in-main-heading searches in UM-D's Blue System. The precision score for the

keyword-in-main-heading searches was quite high but it was obtained from the results of only six searches. Earlham's Blue System submitted no queries to the keyword-in-main-heading search. Generally, precision scores were higher at UM-D (70.0%) than at Earlham (51.1%).

11.2.3 Browsing and Display Opportunities in Controlled Vocabulary Searches

11.2.3.1 Browsing Options

The exact approach gave users several opportunities to browse terminology connected with the subject headings their queries matched or the subject headings they selected from alphabetical lists. Tables 11.2 and 11.3 list browsing and display opportunities in the Blue and Pinstripe Systems and the number of times users selected these options in controlled vocabulary searches at UM-D and Earlham, respectively.

UM-D users conducted 149 exact searches in the Blue System. Most browsing/display opportunities of which users took advantage were displaying titles or explicitly quitting their search. If events for displaying the exact-approach main menu, displaying titles, or quitting searches were not considered in the total number of exact-search opportunities, then searchers hardly took more than one opportunity per search. Despite the availability of broader and narrower terms and subject subdivisions for many subject headings, few users chose these options to continue their searches.

UM-D users conducted 64 alphabetical searches in the Blue System. The browsing/display opportunities they were most likely to choose were selecting terms from or browsing backward or forward in the alphabetical list, expanding their search, authorizing the system to search for a subject heading following the user's entry of a *see* reference, or explicitly quitting the search. Users displayed few broader terms, narrower terms, or subject subdivisions.

Users performed no browsing in connection with subject headings retrieved through the keyword-in-main-heading search.

Table 11.3. Browsing/Display Opportunities (UM-D)

	Blue System			Pinstripe System
Browsing/display opportunities	**Exact**	**Alpha-betical**	**Keyword-in-main**	**Alpha-betical**
Display main menu (exact approach only)	149	n/a	n/a	n/a
Display alphabetical list of subject headings (alphabetical and keyword-in-main searches only)	n/a	64	6	53
Select term or browse forward or backward in alphabetical list (alphabetical and keyword-in-main searches only)	n/a	46	4	60
Choose authorized subject headings in response to entry of *see* reference	41	24	0	25
Display titles	139	27	5	76
Browse broader terms	16	3	0	25
Browse narrower terms	32	4	0	26
Display scope note	3	0	0	5
Display topical subdivisions	24	4	0	13
Display geographic subdivisions	4	0	0	6
Display form subdivisions	14	3	0	11
Display period subdivisions	0	0	0	0
Expand search (Blue System only)	26	31	3	n/a
Redisplay original main menu or alphabetical list	8	4	0	15
Quit search	106	29	2	39
Total	562	239	20	354
Av. no. of opportunities taken per search	3.8	3.7	3.3	6.7

Generally, users conducting controlled vocabulary searches in the Blue System took fewer than four opportunities per search. They took almost seven opportunities per alphabetical search in the Pinstripe System. The browsing/display opportunities in the Pinstripe System they were most likely to choose were selecting terms from or browsing backward or forward in the alphabetical list, expanding their search, displaying titles, or explicitly quitting their search. They were also likely to display broader or narrower terms.

Table 11.4. Browsing/Display Opportunities (Earlham)

Browsing/display opportunities	Blue System		Pinstripe System
	Exact	Alpha-betical	Alpha-betical
Display main menu (exact approach only)	23	n/a	n/a
Display alphabetical list of subject headings (alphabetical approach only)	n/a	12	11
Select term or browse forward or backward in alphabetical list (alphabetical approach only)	n/a	6	17
Choose authorized subject headings in response to entry of *see* reference	7	1	1
Display titles	26	2	11
Browse broader terms	3	0	1
Browse narrower terms	6	1	1
Display scope note	2	0	1
Display topical subdivisions	6	0	5
Display geographic subdivisions	2	0	6
Display form subdivisions	2	0	1
Display period subdivisions	0	0	2
Expand search (Blue System only)	9	7	n/a
Redisplay original main menu or alphabetical list	0	0	4
Quit search	7	6	10
Total	93	35	71
Av. no. of opportunities taken per search	4.0	2.9	6.5

Earlham users conducted almost two dozen (23) exact searches in the Blue System. Most browsing/display opportunities of which users took advantage were displaying the exact-approach main menu, displaying titles, choosing authorized subject headings in response to their entry of *see* references, and expanding their searches with the results of a different subject search approach. Despite the availability of broader and narrower terms and subject subdivisions, few users chose these options to continue their searches.

Earlham users conducted a dozen alphabetical searches in the Blue System. The browsing/display opportunities they were most likely to choose were selecting from or browsing backward or forward in the alphabetical list, expanding their search, or explicitly quitting the search. Users selected few listed terms from alphabetical lists. They also hardly displayed titles, broader or narrower terms, or subject subdivisions.

Earlham users conducted only eleven alphabetical searches in the Pinstripe System. The browsing/display opportunities they were most likely to choose were displaying titles, explicitly quitting the search, and, especially, browsing backward or forward in the alphabetical list. Alphabetical searches in the Pinstripe System averaged between six and seven opportunities per search; Blue-system searches averaged four or fewer opportunities per search.

The amount of browsing UM-D and Earlham users performed in the Pinstripe System was expected because this system's selection of the alphabetical approach was not based on the extent to which listed terms matched user queries; consequently, users performed a lot of alphabetical browsing to find subject headings that described their topics of interest.

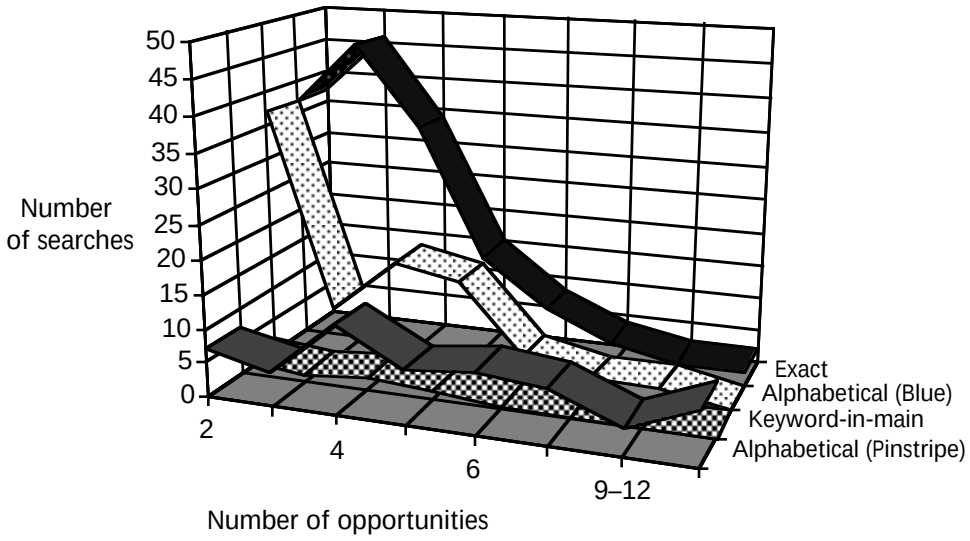

Figure 11.2. Number of opportunities per search (UM-D and Earlham)

Figure 11.2 shows the number of opportunities users took per search. Although users took few opportunities in controlled vocabulary searches, they took even fewer opportunities in Blue-system searches than in Pinstripe-system searches.

Generally, users conducting exact searches took two to four opportunities per search. Only four exact searches exceeded eight opportunities. The longest exact search featured 24 opportunities. This search was initiated by the query “fluids.” The user browsed over a half dozen narrower terms, two broader terms, displayed titles assigned these broader and narrower terms, and displayed topical subdivisions connected with a few of these terms. Generally, users conducting alphabetical searches in the Blue System took two to five opportunities per search. Keyword-in-main searches were also quite short; they ranged from two to five opportunities per search.

Several alphabetical searches in the Pinstripe System were quite long; eight searches featured thirteen or more browsing/display opportunities. Although the Pinstripe System randomly chose the alphabetical search in these eight searches, all eight searches benefited from terms in the alphabetical list that described their topics of interest. The query “composites” initiated the longest search that featured 34 browsing/display opportunities. The user selected the term “Composite materials” from the alphabetical list. The user browsed eight narrower terms, a half dozen broader terms, a few subject subdivisions, and displayed titles assigned to some terms and subdivided subject headings.

Lengthy searches that featured many browsing/display opportunities were the exception rather than the rule in the experimental online catalog’s controlled vocabulary searches. Surprised by the small amount of browsing users did in the Blue and Pinstripe Systems, the researchers performed a failure analysis of successful controlled vocabulary searches to shed light on browsing in the experimental online catalogs. They divided successful controlled vocabulary searches into the following four categories:

- Simple, successful searches — searches in which users performed no browsing to find useful titles. Users only selected the <display general works> option from the exact-approach main menu, and displayed and rated one or more titles.
- Successful searches with a little browsing — searches in which users selected one browsing option from the exact-approach main menu (i.e., notes, narrower terms, broader terms, subject subdivisions) in addition to displaying titles connected with the browsed term or the original matched term.
- Successful searches with more than a little browsing — searches in which users selected two or more browsing options from the exact-approach main menu in addition to displaying titles connected with the browsed term or the original matched term.
- Successful searches due to expanding — searches users conducted in the Blue System in which they expanded the search to find useful titles.

Table 11.5. Browsing in Successful Controlled Vocabulary Searches

	UM-D		Earlham	
Search category	**Blue**	**Pinstripe**	**Blue**	**Pinstripe**
Simple search				
Exact	55	n/a	7	n/a
Alphabetical	14	10	0	2
Keyword-in-main	3	n/a	0	n/a
Subtotal	72 (59.5%)	10 (37.1%)	7 (50.0%)	2 (40.0%)
A little browsing				
Exact	12	n/a	4	n/a
Alphabetical	1	9	0	3
Keyword-in-main	0	n/a	0	n/a
Subtotal	13 (10.8%)	9 (33.3%)	4 (28.6%)	3 (60.0%)
More than a little browsing				
Exact	12	n/a	1	n/a
Alphabetical	2	8	0	0
Keyword-in-main	1	n/a	0	n/a
Subtotal	15 (12.4%)	8 (29.6%)	1 (7.1%)	0 (0.0%)
Expanded search				
Exact	5	n/a	1	n/a
Alphabetical	15	n/a	1	n/a
Keyword-in-main	1	n/a	0	n/a
Subtotal	21 (17.3%)	n/a	2 (14.3%)	n/a
Total	121 (100%)	27 (100%)	14 (100%)	5 (100%)

These four successful-search categories are listed in Table 11.5. Each category is further subdivided by the particular subject searching approach that the Blue or Pinstripe System invoked to find a match.

Overall, simple searches made up over half of successful searches. Expanded searches were somewhat common in the Blue System where they accounted for about one of every eight searches.

Sections 11.2.3.2–11.2.3.6 describe the failure analysis of browsing in successful controlled vocabulary searches.

11.2.3.2 Simple, Successful Controlled Vocabulary Searches

In simple, successful controlled vocabulary searches, users performed almost no browsing to find useful titles. Of the 167 successful controlled vocabulary searches, 91 (54.5%) were simple, successful controlled vocabulary searches.

Table 11.6 describes simple, successful controlled vocabulary searches in the Blue and Pinstripe Systems.

Table 11.6. Simple, Successful Controlled Vocabulary Searches

User query	Explanation
cad	Blue System informs the user that the entered term is a cross reference to "Computer-aided design." It responds with the results of an exact search in which options for displaying general works, narrower terms, subject and form subdivisions, and for expanding the search are enabled. User displays 2 of 22 retrieved titles and rates both "very useful."
cast iron	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, subject subdivisions, and for expanding the search are enabled. User displays all 5 retrieved titles and rates them "very useful."
american indian religious freedom act	Pinstripe System responds with the alphabetical approach in which "American Indians" is the closest matching term. User selects this *see* reference for "Indians of North America," retrieves 12 titles, and rates one title useful.
c language	Blue System responds with results of a keyword-in-main-heading search in which three subject headings contain the same words as the user query (i.e., "C (Computer program language)," "PL/C (Computer program language)," and "Small-C (Computer program language)." User selects the former, displays all 42 retrieved titles, and rates all but one title "very useful."

11.2.3.3 Successful Controlled Vocabulary Searches with a Little Browsing

Successful controlled vocabulary searches that were characterized by "a little browsing" were searches in which users selected one browsing option from the exact-approach main menu (i.e., notes, narrower terms, broader terms, subject subdivisions) in addition to displaying titles connected with the browsed term or the original matched term or expanding the search.

Table 11.7 describes successful controlled vocabulary searches with a little browsing.

Table 11.7. Successful Controlled Vocabulary Searches With a Little Browsing

User query	Explanation
atomic bomb	Blue System responds with the results of an exact search in which options for displaying general works, broader terms, topical subdivisions, and for expanding the search are enabled. User browses topical subdivisions, selects "History," displays one title and rates it "very useful."
operating systems	Blue System responds with the results of an exact search in which options for displaying general works, narrower terms, topical and form subdivisions, and for expanding the search are enabled. User retrieves 32 titles for "Operating systems," displays 17 titles, and rates 5 titles "very useful." User displays narrower terms, chooses "Distributed operating systems (Computers)," and rates the 1 retrieved title (*Open systems)* "very useful."
women in the westward movement	Pinstripe System responds with the alphabetical approach in which "Women pioneers" is the closest matching term. User selects this heading, chooses geographical subdivisions connected with the heading, and retrieves three titles which she rates "very useful."
heat exchangers design	Pinstripe System responds with the alphabetical approach in which "Heat insulating materials" is the closest matching term. User chooses preceding term "Heat exchangers." Exact options for displaying titles, browsing broader or narrower terms, topical and form subdivisions, and expanding the search are enabled. User displays the 13 retrieved titles and gives them negative ratings. He chooses topical subdivisions option, selects "Design and construction," and rates the two retrieved titles (*Industrial heat exchangers* and *Heat exchangers: selection, design, and construction*) "very useful."

Of the four successful searches described in Table 11.7, two searches were conducted in the Pinstripe System. Although the terminology of user queries and displayed subject headings differed, users found suitable expressions for their topics of interest by browsing the alphabetical list of subject headings and/or by browsing subdivided forms of selected subject headings.

11.2.3.4 Successful Controlled Vocabulary Searches with More than a Little Browsing

Successful controlled vocabulary searches that were characterized by "more than a little browsing" were searches in which users selected more than one browsing option

from the exact-approach main menu (i.e., notes, narrower terms, broader terms, subject subdivisions) in addition to displaying titles connected with the browsed term or the original matched term or expanding the search.

Three successful controlled vocabulary searches with more than a little browsing are described in Table 11.8.

Table 11.8. Successful Controlled Vocabulary Searches With More than a Little Browsing

User query	Explanation
medical	Pinstripe System responds with the results of an alphabetical search in which "Medical care — Moral and ethical aspects" is the closest matching term. During this lengthy search, the user chooses several terms from the alphabetical list, e.g., "Medical statistics," "Medical writing," "Medicine," displays broader terms, narrower terms, notes, subject subdivisions, and titles.
heat transfer	Blue system informs the user that the entered term is a cross reference to "Heat — Transmission." It responds with the results of an exact search. User displays 42 titles assigned to the subject heading and gives all titles negative ratings. User displays narrower terms, selects "Heat exchangers" from a list of five terms, displays 13 titles, rates several useful. User also finds useful records under this subject heading that are subdivided by "Design and construction."
camera	Blue System responds with the results of an alphabetical search in which "Camera lenses," "Camera operators, Motion pictures," and "Cameras" begin with the same word as the user query. User chooses the former which is a *see* reference to "Lenses, Photographic," rates all 7 retrieved titles "very useful." User selects more terms from alphabetical list, e.g., "Photography," "Cameras," and rates titles useful.

11.2.3.5 Successful Controlled Vocabulary Searches with Expanding

Successful controlled vocabulary searches that were characterized by expanding were searches users conducted in the Blue System in which they expanded the search to find useful titles.

Table 11.9 describes four successful controlled vocabulary searches with expanding.

Table 11.9. Successful Controlled Vocabulary Searches With Expanding

User query	Explanation
psychology	Blue System responds with the results of an alphabetical search in which "Psychology" is the closest matching term. User selects this heading and the system displays enabled exact-search options for displaying narrower terms, topical subdivisions, and for expanding the search. User chooses the expand option and the system retrieves 18 titles from a title-keyword search. User displays 18 titles and gives 4 titles "very useful" ratings.
mechanical engineering	Blue system informs the user that the entered term is a cross reference to "Mechanics, Applied." In response to exact-search main menu options, user selects expand option and system responds with the results of an alphabetical search. User chooses "Mechanics, Dynamic" from the alphabetical list, retrieves one title, and gives it a "very useful" rating.
flexible manufacturing systems	Blue System responds with the results of an exact search. In response to exact-search main menu options, user selects expand option and system responds with the results of an alphabetical search. User chooses "Flexible manufacturing systems" from the alphabetical list, retrieves 14 titles, and gives them a "very useful" rating.

Since many experimental online catalog searchers were using the system for the first time, some searchers might have thought that selecting the Blue System's "expand" option was the way to display titles connected with a selected subject heading. For example, in the "flexible manufacturing systems" search, the searcher disregarded the "display general works" option, selected the "expand this search" option, and completed the alphabetical search with search options that would have been the same for completing the exact search for the same subject.

Future implementations of an expand capability should feature experimentation with users to ensure that understandable terminology is used to indicate options for displaying titles and expanding the search with the results of a different subject searching approach.

11.2.3.6 Discussion of Browsing in Successful Controlled Vocabulary Searches

Despite the browsing opportunities that controlled vocabulary searches offered, end users performed little browsing. Users seemed satisfied with displaying titles under the "general works" option and terminating the search. More browsing characterized alphabetical searches, especially in the Pinstripe System. The Pinstripe System

responded randomly with the alphabetical approach regardless of the extent of match between user queries and subject headings. (The Blue System responded with the alphabetical approach when user queries matched the initial words of subject headings in the alphabetical list.) Perhaps users sensed that their terms were not exact matches of subject headings and were predisposed to browse using enabled opportunities. Interestingly, the lengthy browsing that characterized a handful of Pinstripe-system searches was initiated by queries that were exact or almost exact matches of controlled vocabulary terms.

When users initiated searches that matched controlled vocabulary terms, they were unlikely to browse for additional terminology or refine their original terminology with more specific terminology that narrower terms and subject subdivisions could have provided. Perhaps the searchers that really need browsing opportunities are those whose terms are not close to or controlled vocabulary terms. In the absence of post-search interview data that explores browsing in subject searches, we can only speculate on the paucity of browsing in controlled vocabulary searches. This may also be a fruitful area for additional research.

11.2.4 Perseverance in Controlled Vocabulary Searches

The failure analysis in chapter 10 featured discussions of perseverance in subject searches (sections 10.2.1.1, 10.2.2.1, 10.3.1.1, and 10.3.2.1). Perseverance was a major problem in controlled vocabulary searches. Of the 272 controlled vocabulary searches in the Blue and Pinstripe Systems at UM-D, retrieval in 63 (23.2%) searches was marred by user perseverance. Of the 46 controlled vocabulary searches in the Blue and Pinstripe Systems at Earlham, retrieval in eight (15.4%) searches was marred by user perseverance.

Table 11.10 describes four searches that were adversely affected by user perseverance. These four searches were typical examples of controlled vocabulary searches in the Blue and Pinstripe Systems in which the systems were successful in matching user queries with subject headings that described users' topics of interest. Unfortunately, users did not display titles assigned those subject headings.

Table 11.10. Controlled Vocabulary Searches Marred by User Perseverance

User query	Explanation
turbo prolog	Blue System responds with the results of an exact search in which exact-approach main menu options for displaying general works, broader terms, and expanding the search are enabled. User quits. Had she selected the "general works" option, she would have retrieved one title, *Introduction to Turbo prolog.*
underground railroad	Blue System responds with the results of an exact search in which exact-approach main menu options for displaying general works, broader terms, notes, geographic subdivisions, and expanding the search are enabled. User quits. Had he selected the "general works" option, he would have retrieved several titles on this topic, e.g., *The Underground Railroad: from slavery to freedom, The liberty line: the legend of the Underground Railroad.*
automatic transmission	Blue system responds with the results of an alphabetical search in which "Automatic transmissions, Automobile" is the closest match. The user selects it and the system informs the user that the selected term is a cross reference to "Automobiles — Transmission devices, Automatic." The system responds with the results of an exact search with options for displaying general works and expanding the search. User switches systems and misses potentially useful titles such as *Automatic transmission* and *Automatic and fluid transmissions.*
c programming language	Pinstripe system responds with the results of an alphabetical search in which "C (Computer program language)" is the closest match. The user selects it and the system responds with the results of an exact search with options for displaying general works, broader and narrower terms, and expanding the search. User switches systems and misses several potentially useful titles.

In the absence of post-search interview data that probed for reasons why users terminated these seemingly successful searches, we can speculate on such reasons. One reason may be the density or complexity of the exact-approach main menu on which was listed options for displaying titles and browsing subject terms. Figure 11.3 shows the exact-approach main menu that the Blue System displayed to users in exact searches and exact-search components of alphabetical and keyword-in-main-heading searches.

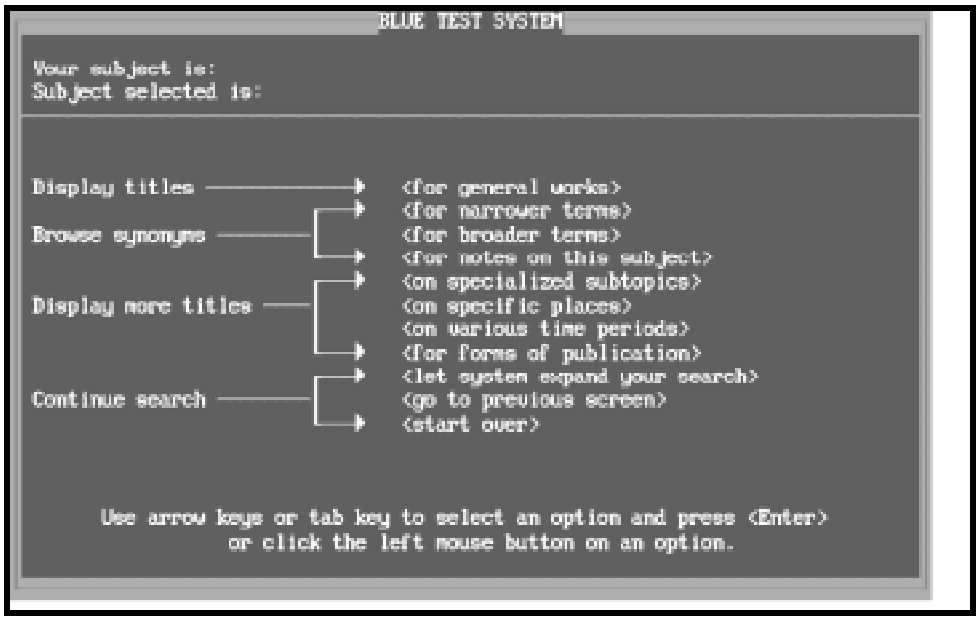

Figure 11.3. Exact-approach main menu

The researchers were limited in the design of the exact-approach main menu by the FoxPro database management system. After considerable discussion amongst project team members and consultants, we accepted the design of the menu in figure 11.3 in which phrases on the left side of the menu described possible next steps in the search. These phrases pointed to specific options on the right side of the menu. Options were enabled or disabled depending on the particular matched subject heading. They were also "hot" in that selection of a particular option provoked the system to perform the described action.

The menu was dense and complex. The screen contained mostly text with little in the way of graphics except for arrows leading from phrases to specific options. For some users, this menu could have been so dense or complex that they chose to start over rather than study or experiment with listed options. Other users might not have understood that the "display general works" option would have resulted in a display of retrieved titles for the matched subject heading. Section 11.2.3.5 described several searches in which the logical next step would have been for users to choose the "display general works" option, but, instead, users chose the "expand search" option. Perhaps they did not understand the meaning of the "display general works" option and/or "expand search" option.

The design of the exact-approach main menu is critical to the exact search approach's success in retrieving titles on users' topics of interest. Future implementations of the

exact search must include end users in the design and pretesting of this menu, especially in the wording of listed options.

11.2.5 Navigation Problems in Controlled Vocabulary Searches

The failure analysis in chapter 11 featured a general discussion of navigation in subject searches (sections 10.2.1.8, 10.2.2.8, and 10.3.2.8). Navigation was not an obvious problem in controlled vocabulary searches because we did not conduct post-search interviews with end users to probe reasons for search failures. The researchers inferred navigation problems based on the illogical or strange actions users took in the course of their searches. For example, in response to the user-entered query "plastic molding," the Pinstripe System responded with the alphabetical approach in which the term "Plastic molding" was the closest match of the query. When the user selected this term, the system informed him that it was a *see* reference for "Plastics — Molding" and responded with an exact-approach main menu with enabled options for displaying titles, broader and narrower terms, form subdivisions, and expanding the search. The user browsed narrower terms, selected "Injection molding of plastics" and displayed subdivided forms of this subject heading. Instead of displaying the original alphabetical list bearing the term that described his query, i.e., "Plastic molding" and displaying titles on this topic, the user terminated the search. Most likely, the user did not understand how to backtrack in the search and terminated the search with the intention of starting over.

Of the 272 controlled vocabulary searches in the Blue and Pinstripe Systems at UM-Dearborn, 15 (5.5%) searches exhibited likely navigation problems. Of the 46 controlled vocabulary searches in the Blue and Pinstripe Systems at Earlham, six (13.0%) searches exhibited likely navigation problems. Although these are small percentages, many of the searches marred by perseverance might have actually been marred by navigation problems.

Table 11.11 describes four searches in which navigation problems contributed to the users' failure in retrieving useful titles.

Table 11.11. Navigation Problems in Controlled Vocabulary Searches

User query	Explanation
common business oriented language	Blue system informs the user that the entered term is a cross reference for "COBOL (Computer program language)." In response to exact-search main menu options, user selects broader-term option and selects the listed broader term "English language." User browses form subdivisions and chooses "Dictionaries." User browses topical subdivisions, e.g., "Arabic," "French," "German," and quits searching. The user's selection of the broader term "English language" dealt the killing blow to this search because it lead him away from his topic of interest.
animation	Blue system responds with the results of an alphabetical search in which "Animation, Computer" is the closest match. User browses forward in the alphabetical list, selects "Casting, Steel" which is a *see* reference for "Steel founding." User quits the search.
african-americans	Blue system informs the user that the entered term is a cross reference for "Afro-Americans." In response to exact-search main menu options, user selects narrower-term option and selects the narrower term "United States — Civil War, 1861–1865 — Afro-American influences." User quits the search. Most likely, user wants to backtrack to the exact-search main menu to explore other terms or display titles but does not know what options to pick to return to this menu.
civil rights	Pinstripe system responds with the results of an alphabetical search in which "Civil rights" is the closest match. In response to exact-search main menu options, user browses geographic subdivisions, chooses "Florida." In response to exact-search main menu options, user browses geographic subdivisions, chooses "Saint Augustine." In response to exact-search main menu options, user browses topical subdivisions, chooses "History." In response to exact-search main menu options, user browses geographic subdivisions, chooses "20th century." In the post-search interview that followed this librarian's search, he comments on the inefficiency of intervening exact-search main menus.

The navigation problems described in this section call for the following improvements to the exact search:

- Reducing the complexity of the exact-approach main menu.
- Consolidating several exact-approach main menu options into fewer options based on numbers of retrievals.

- Including end users in the design and pretesting of the exact-approach main menu.
- Pretesting the wording of exact-approach main menu options with end users.
- Designing and implementing an easy, understandable backtracking option.

11.2.6 Alphabetical Approach as a Last Resort

Although the Pinstripe System's selection of the alphabetical search was done by a random selection algorithm, there were several searches in which the closest matching term or nearby terms in the alphabetical list were not close matches in terms of syntax but users eventually found useful titles by selecting them. Several examples are given in Table 11.12.

Table 11.12. Queries and Matching Terms in Pinstripe Alphabetical Searches

User queries	Close matches in meaning
american indian religious freedom act	"American Indians" (*see* reference for "Indians of North America")
women in the westward movement	"Women pioneers"
florida key	"Florida"
computers made simple	"Computers"
emissions	System placement at "Employee incentives;" user browses forward and finds "Emission control devices" (*see* reference for "Pollution control devices").
heat exchangers design	System placement at "Heat insulating materials;" user browses forward and backward, finds "Heat exchangers," and, eventually, finds subdivision "Design and construction."

In view of users' experiences with the Pinstripe System's alphabetical approach, the researchers recommend this approach as a last resort. That is, when user queries fail to meet the criteria for controlled vocabulary or free-text approaches, search trees should force systems to respond with the alphabetical approach regardless of the extent of match between user queries and subject terms in the alphabetical list. On occasion, listed subject terms may describe the topics users have in mind and provide them with the opportunity to browse forward or backward in the list to find suitable terms.

11.3 Keyword-in-Subdivided-Heading Search

11.3.1 Review of the Keyword-in-Subdivided-Heading Search

Both Blue and Pinstripe Systems featured the keyword-in-subdivided-heading search. Like the exact, alphabetical, and keyword-in-main-heading search, this search was a controlled vocabulary search. Unlike the exact, alphabetical, and keyword-in-main-heading search, the user's entry of a subject query or selection of a system-displayed subject heading did not result in the exact approach. Instead, the systems responded with an alphabetical list of main and subdivided subject headings bearing the same words as those in user queries. The user's selection of a listed subject heading resulted in a display of titles assigned the selected subject heading. When users terminated the display of titles, the systems returned to the display of matching subject headings so that users could select more subject headings of interest and display more titles.

The Blue and Pinstripe Systems responded to user queries with the keyword-in-subdivided-heading search under different circumstances. The Blue System responded with keyword-in-subdivided-heading search results when the words in user-entered queries matched subject headings bearing the same words, and, usually, additional words. For example, the user query "afro americans and west" invoked the keyword-in-subdivided-heading search in the Blue System. It responded with two subdivided subject headings: "Afro-Americans — West (U.S.)" and "Afro-Americans — Middle West." Both subject headings contained the same words as the words in user queries, i.e., "afro," "americans," and "west," and additional words, i.e., "us" and "middle."

A random selection algorithm governed the Pinstripe System's selection of the keyword-in-subdivided-heading search. The Pinstripe System responded with this search regardless of the particular elements in user queries.

Table 11.13 summarizes the reasons for success or failure of keyword-in-subdivided-heading searches.

Table 11.13. Keyword-in-subdivided-heading Searches

Reasons for success/failure	UM-D		Earlham	
	Blue	Pinstripe	Blue	Pinstripe
Successful searches:				
Simple search	1	25	2	1
A little browsing	0	9	0	3
More than a little browsing	0	1	0	0
Expanded search	1	n/a	1	n/a
Subtotal	2 (100.0%)	35 (28.2%)	3 (42.9%)	4 (17.4%)
Unsuccessful searches:				
Perseverance	0	24	0	0
No matches	0	28	0	12
Database failure	0	23	2	5
Query specificity	0	5	2	0
Too many retrievals	0	4	0	0
Search approach failures	0	5	0	0
Vocabulary	0	0	0	2
Subtotal	0 (0.0%)	89 (71.8%)	4 (57.1%)	19 (82.6%)
Total	2 (100%)	124 (100%)	7 (100%)	23 (100%)

Keyword-in-subdivided-heading searches were much more frequent in the Pinstripe System than in the Blue System. We attributed the search trees as the reason for this. A random selection algorithm governed the Pinstripe System's selection of the keyword-in-subdivided-heading search. The random selection algorithm had three subject searching approaches from which to choose. Thus, the Pinstripe System should have responded to about a third of user-entered queries with the keyword-in-subdivided-heading search. Search trees governed the Blue System's selection of a particular subject searching approach. They favored exact, alphabetical, and keyword-in-main-heading searches before the keyword-in-subdivided-heading search. Although users frequently entered queries into the Blue System that matched words in subject headings, this system responded to many of these queries with the exact, alphabetical, or keyword-in-main-heading searches. Blue-system users could still obtain the results of keyword-in-subdivided-heading searches for such queries by repeatedly expanding their searches.

Overall, unsuccessful keyword-in-subdivided-heading searches outnumbered successful keyword-in-subdivided-heading searches. Almost three-quarters (112 of 156) of keyword-in-subdivided-heading searches were unsuccessful. Frequent reasons for failure

were perseverance, database failure, and keyword-in-subdivided-heading searches that retrieved no subject headings in the Pinstripe System.

This section describes keyword-in-subdivided-heading searches in detail, focuses on strengths of this search, and makes suggestions for improving its weaknesses.

11.3.2 Precision in Keyword-in-subdivided-heading Searches

Figure 11.4 shows precision scores for keyword-in-subdivided-heading searches. (Precision scores included conflicting relevance assessments and relevance assessments for repeated displays of the same titles.)

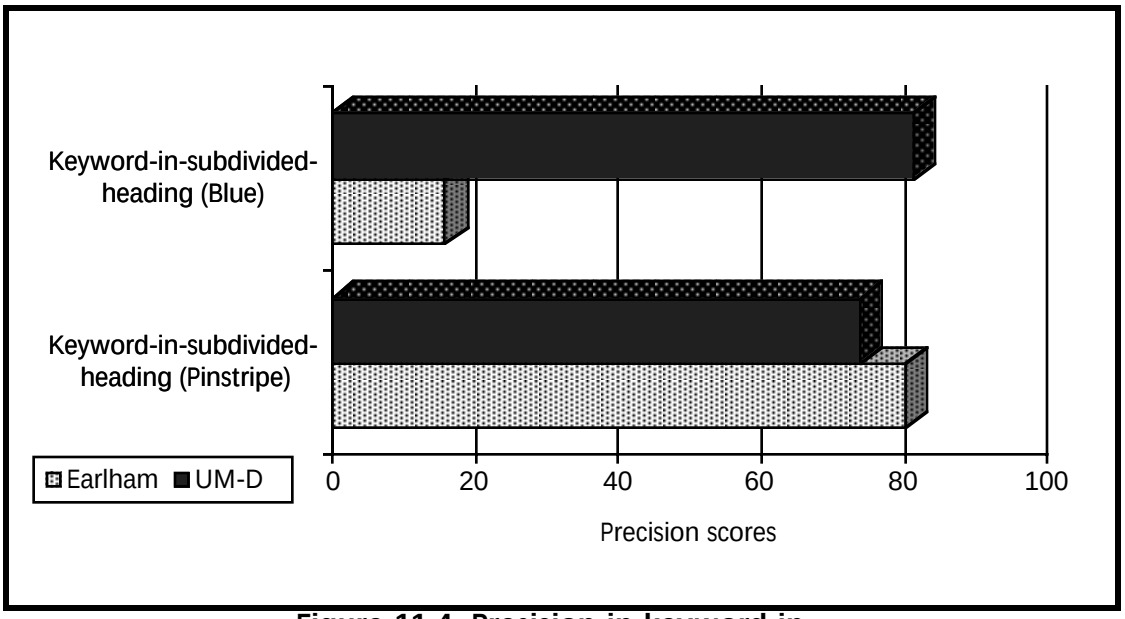

Figure 11.4. Precision in keyword-in-subdivided-heading searches

Except for precision in Earlham's Blue System, precision scores for keyword-in-subdivided-headings searches were quite high. Precision scores in Earlham's Blue System were based on the results of only seven searches.

11.3.3 Successful Keyword-in-subdivided-heading Searches

Table 11.13 lists four categories of successful keyword-in-subdivided-heading searches. Descriptions of these four categories are:

- Simple searches — searches in which users selected one listed subject heading, displayed one or more titles, and gave "very useful" ratings to one or more titles.
- A little browsing — searches in which users selected two listed subject headings, displayed one or more titles under each heading, and gave "very useful" ratings to one or more titles in the course of the search.
- More than a little browsing — searches in which users selected more than two listed subject headings, displayed one or more titles under more than two subject headings, and gave "very useful" ratings to one or more titles in the course of the search.
- Successful searches due to expanding — searches users conducted in the Blue System in which they expanded the search to find useful titles.

Successful keyword-in-subdivided-heading searches numbered 44. The average number of subject headings retrieved in these successful searches was 10.6. The range of retrieved subject headings was from 1 to 121. If postings for the one search that resulted in 121 subject headings were disregarded in the calculation of the average number of retrieved subject headings, this average would drop to eight subject headings per successful search. Nine keyword-in-subdivided-heading searches retrieved only one subject heading.

Two-thirds (29 of 44) of successful keyword-in-subdivided-heading searches featured no browsing. These were simple searches in which users selected one listed subject heading, displayed one or more titles, and gave "very useful" ratings to one or more titles. Only one of 44 successful keyword-in-subdivided-heading searches featured more than a little browsing. Table 11.14 gives examples of the four types of successful keyword-in-subdivided-heading searches.

Users performed little browsing in keyword-in-subdivided heading searches. The design of this search encouraged browsing because it always returned to the list of retrieved subject headings following the display of retrieved titles. Instead of choosing another subject heading, most users chose to terminate the search.

Table 11.14. Successful Keyword-in-subdivided-heading Searches

User query	Explanation
civil war historiography	Simple search in which the Blue System retrieves one subdivided subject heading, "United States — History — Civil War, 1861–1865 — Historiography." User selects this heading, retrieves six titles, displays all six, and rates one "very useful."
inventions	Search that features "a little browsing" in which Pinstripe System retrieves a total of seven subject headings (i.e., "Inventions," and six subdivided forms of "Inventions"). The user chooses "Inventions — Congresses" and "Inventions," retrieves 13 titles, and displays one title under each heading, and rates one title "very useful."
polymers	In the one search that features "more than a little browsing," Pinstripe System retrieves a total of 12 subject headings. The user chooses "Polymers," and five subdivided forms of this heading, e.g., "Handbooks, manuals, etc.," "Analysis," "Information services," and rates most displayed titles "very useful."
numerical control	Blue System retrieves 4 subject headings. User chooses "Machine design — Numerical control," retrieves 24 titles, and displays a few. User expands search, retrieves 32 titles in a title-keyword search, and displays some.

11.3.4 Unsuccessful Keyword-in-subdivided-heading Searches

There were many more unsuccessful keyword-in-subdivided-heading searches than successful ones. Of the seven reasons for failure, three reasons accounted for over four-fifths of failed searches: (1) searches that retrieved no subject headings in the Pinstripe System (35.7% of failures), database failure (26.8% of failures), and perseverance (21.4% of failures). This section focuses on these and other reasons for unsuccessful keyword-in-subdivided heading searches.

11.3.4.1 No Postings in Pinstripe-system Searches

The most frequent reason for failure of the keyword-in-subdivided-heading search was the failure of this search to retrieve subject headings in the Pinstripe System. Examples of user queries that failed to retrieve subject headings are:

- american events
- civil rights movement
- abolitionism

- flow charting
- software management
- computing dictionary

Although the Pinstripe System's keyword-in-subdivided-heading search failed to retrieve subject headings for these queries, subject searching approaches in the Blue System were successful retrieving subject headings and useful titles. For example, the Blue System retrieved two titles, *Dictionary of computing and new information technology* and *Dictionary of computing*, that the user rated "very useful" in a title-keyword search.

Table 11.15. Blue-system Search Outcomes for Failed Keyword-in-subdivided-heading Searches

Outcome	UM-D	Earlham
Successful search outcomes		
Title-keyword	8	6
Matching *see* reference in exact search	4	0
Keyword-in-record	2	1
Keyword in subject heading fields	0	1
Exact approach	0	1
Alphabetical approach	2	0
Unsuccessful search outcomes		
Perseverance	2	2
Query specificity	4	1
No matches (search trees exhausted)	2	0
Vocabulary problems	4	0
Total	28	12

Table 11.15 lists the subject searching approaches that the Blue System enlisted for keyword-in-subdivided-heading searches that failed in the Pinstripe System. It also gives the outcome of Blue-system searches.

A lot of keyword-in-subdivided-heading searches that failed in the Pinstripe System were successful in the Blue System through title-keyword searches. A few failed searches were successful in the Blue System because the user query matched a *see* reference in an exact search. Examples were the user queries "database management" and "ergonomics" that were *see* references for "Data base management" and "Human

engineering," respectively. The availability of *see* references for these queries made it possible for users to retrieve titles on their topics of interest.

Other keyword searches figured into the success of a few searches in the Blue System, i.e., keyword-in-record and keyword in subject heading fields searches. Two user queries, "medical statistic" and "power tool" matched the initial characters in longer subject headings, "Medical statistics" and "Power tools," respectively. These two queries triggered the Blue System's alphabetical approach and, eventually, resulted in the retrieval of useful titles.

Generally, the success of searches that failed to produce retrievals through the Pinstripe System's keyword-in-subdivided-heading approach usually succeeded through the Blue System's free-text approaches, i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches, or *see* references featured in controlled vocabulary searches.

11.3.4.2 Database Failure

The failure of the keyword-in-subdivided-heading approach was frequently attributed to database failure. Examples of queries that failed because of database failure were:

- portable data file
- radar detectors
- computer workstation
- women in reconstruction
- dayton history
- minorities in the west

Database failure was to blame for the failure of users' corresponding searches of the Blue System for the same topic for many of these queries.

11.3.4.3 Perseverance

Perseverance was a problem with keyword-in-subdivided-heading searches, especially in UM-D's Pinstripe System where the failure of two dozen searches was attributed to perseverance. Perseverance could not be blamed on high numbers of retrievals because retrievals were not much different from retrievals for successful searches. For keyword-in-subdivided-heading searches marred by perseverance the average number of subject headings retrieved per query was 13.67. The range of retrieved subject headings was from 1 to 144. If postings for the one search that retrieved 144 subject headings were

disregarded in the calculation of the average number of displayed subject headings, this average would drop to eight subject headings per search. This average was the same as the average for successful controlled vocabulary searches minus the one search with the highest postings. Two keyword-in-subdivided-heading searches retrieved one subject heading.

Table 11.16 describes keyword-in-subdivided-heading searches that failed due to user perseverance. This table describes seemingly successful searches. Retrievals were quite manageable in that users did not have to browse more than two screens of retrieved subject headings to find one(s) that interested them. Unfortunately, instead of browsing or selecting listed subject headings, users terminated their searches. Perhaps the low level of perseverance in keyword-in-subdivided-heading searches was indicative of the overall low level of perseverance in controlled vocabulary searches (see section 11.2.4) and low levels of browsing in controlled vocabulary searches (see section 11.2.3) and in keyword-in-subdivided-heading searches (see section 11.3.3).

Table 11.16. Perseverance in Keyword-in-subdivided-heading Searches

User query	Explanation
error correcting	Pinstripe System retrieves 1 subdivided subject heading, "Error-correcting codes." User terminates the search.
robotics	Pinstripe System retrieves 12 subject headings, "Robotics" and 11 subdivided forms of this heading, e.g., "Research," "Mathematics," "Economic aspects." User terminates the search.
composite materials	Pinstripe System retrieves 9 subject headings, "Composite materials" and 8 subdivided forms of this heading, e.g., "Fatique," "Fracture," "Testing." User terminates the search.
computer science	Pinstripe System retrieves 15 subject headings, "Computer science," and several headings with this phrase as a qualifier, e.g., "Natural language processing (Computer science)," "Neural networks (Computer science)," "Expert systems (Computer science)." User terminates the search.

11.3.4.4 Large Retrievals

Of the remaining reasons for failure, large retrievals were especially interesting. The failure of only four searches was attributed to large retrievals. User queries and retrievals in these searches were:

- Mathematics (52 subject headings)
- Computers (117 subject headings)

- Computer (over 400 subject headings)
- Materials (121 subject headings)

The experimental online catalogs did not count up retrieved subject headings and report the number to users. In searches for these four queries, users might have browsed forward and/or backward in the subject headings lists. When they realized how lengthy subject heading displays were, they terminated their searches rather than select listed subject headings.

11.4 Title-keyword Searches

Title-keyword searches were unique to the Blue System. The Blue System's search trees submitted one-word queries to title-keyword searches following unsuccessful exact and alphabetical searches or in response to the user's selection of the expand option following successful exact and/or alphabetical searches. It submitted queries exceeding one word to title-keyword searches following unsuccessful exact, alphabetical, keyword-in-main-heading, and keyword-in-subdivided-heading searches or in response to the user's selection of the expand option following successful exact, alphabetical, keyword-in-main-heading, and/or keyword-in-subdivided-heading searches. The Blue System retrieved titles that contained the same words in user queries and displayed the first of one or more retrieved records to searchers.

Table 11.17 summarizes the number of title-keyword searches and reasons for success or failure.

Table 11.17. Title-keyword Searches

Reasons for success/failure	UM-D	Earlham
	Blue	Blue
Successful searches:		
Simple search	28	9
Expanded search	8	4
Subtotal	36 (69.2%)	13 (86.7%)
Unsuccessful searches:		
Perseverance	8	1
Database failure	5	0
Query specificity	3	1
Subtotal	16 (30.8%)	2 (13.3%)
Total	52 (100%)	15 (100%)

Title-keyword searches were not very common in the Blue System. This was due to the several searches preceding title-keyword searches (i.e., exact, alphabetical, keyword-in-main-heading, and keyword-in-subdivided-heading) to which the search trees submitted user queries.

Precision was not as high in title-keyword searches as in controlled vocabulary searches. Precision scores for title-keyword searches were 47.5% and 23.3% at UM-D and Earlham, respectively.

Successful title-keyword searches outnumbered unsuccessful title-keyword searches. Almost 70% of title-keyword searches in UM-D's Blue System were successful; over 85% of title-keyword searches in Earlham's Blue System were successful. Title-keyword searches were low posted. They averaged 6.3 titles and ranged from zero to 68 titles.

On occasion, users rated several titles "very useful" that contained the same main or subdivided subject heading. A system that features relevance feedback could automatically examine subject headings in useful titles. Users who want additional material on their topics of interest could benefit from the system's retrieval of titles bearing these subject headings. Table 11.18 describes several title-keyword searches with such headings.

Table 11.18. Frequently-occurring Subject Headings in Title-keyword Searches

User query	Explanation
vietnam war	Blue System retrieves 8 titles. Of the 5 rated "very useful" by the user, most bear subdivided forms of the main subject heading, "Vietnamese Conflict, 1961–1975."
american events	Blue System retrieves 7 titles. The 1 title rated "very useful" bears the subdivided subject heading "United States — History — Chronology." Examples of titles retrieved with this subject heading are *The encyclopedia of American facts and dates* and *Webster's guide to American history: a chronological, geographical, and biographical survey.*
computing dictionary	Blue System retrieves 2 titles. 1 of the 2 titles rated "very useful" bears the subdivided subject heading "Computers — Dictionaries." Examples of additional titles retrieved with this subject heading are *The computer dictionary, Computer dictionary*, and *McGraw-Hill dictionary of electronics and computers.*
turbomachinery	Blue System retrieves 8 titles. The 3 titles rated "very useful" bear the subject heading "Turbomachines." A title retrieved with this subject heading is *Theory of turbomachines.*

Perseverance was not as much of a problem in title-keyword searches as it was in controlled vocabulary searches. Eight title-keyword searches were marred by perseverance. Despite the low postings (one or two titles) in four of the eight searches, users walked away from these searches without giving relevance assessments to retrieved titles.

The failure of five title-keyword searches was attributed to database failure. Most of these searches retrieved a few titles but they were false drops. An example was the title-keyword search for "workstation" that retrieved one title entitled *Workstation environment for wastewater treatment design*.

11.5 Keyword in Subject Heading Field Searches

Only the Blue System featured keyword in subject heading field searches. The Blue System submitted queries exceeding one word to keyword in subject heading field searches following unsuccessful exact, alphabetical, keyword-in-main-heading, keyword-in-subdivided-heading searches, and title-keyword searches or in response to the user's selection of the expand option following successful controlled vocabulary and/or other free-text searches. There were only four keyword in subject heading field searches. Two searches each were conducted in Earlham's catalog and in UM-D's

catalog. The small numbers of keyword in subject heading field searches was due to the several searches preceding this search to which the search trees submitted user queries.

Three of the four searches were simple, successful searches in which users retrieved a few titles and rated them "very useful." The fourth search failed because of a vocabulary problem. The query "parallel programming languages" retrieved one title that the user rated "possibly useful." If the "languages" element of the query were deleted from this query, the Blue System could have retrieved several additional titles on the user's topic of interest.

Precision scores for the four searches were not very high. The two searches conducted in UM-D's catalog averaged 50.0% precision and the two searches conducted in Earlham's catalog averaged 35.3% precision.

11.6 Keyword-in-record Searches

11.6.1 Review of Keyword-in-record Searches

Both Blue and Pinstripe Systems featured keyword-in-record searches. The Blue System submitted queries exceeding one word to keyword-in-record searches as a last resort, that is, following unsuccessful exact, alphabetical, keyword-in-main-heading, keyword-in-subdivided-heading searches, title-keyword, and keyword in subject heading fields searches or in response to the user's selection of the expand option following successful search approaches. A random selection algorithm governed the Pinstripe System's selection of the keyword-in-subdivided-heading search. The Pinstripe System responded with this search regardless of the particular elements in user queries.

The two systems retrieved titles that contained the same words in title, subject heading, and other subject-bearing fields of bibliographic records as in user queries and displayed the first of one or more retrieved records to searchers.

11.6.2 Precision in Keyword-in-record Searches

Figure 11.5 shows precision scores for keyword-in-record searches. (Precision scores include conflicting relevance assessments and relevance assessments for repeated displays of the same titles.)

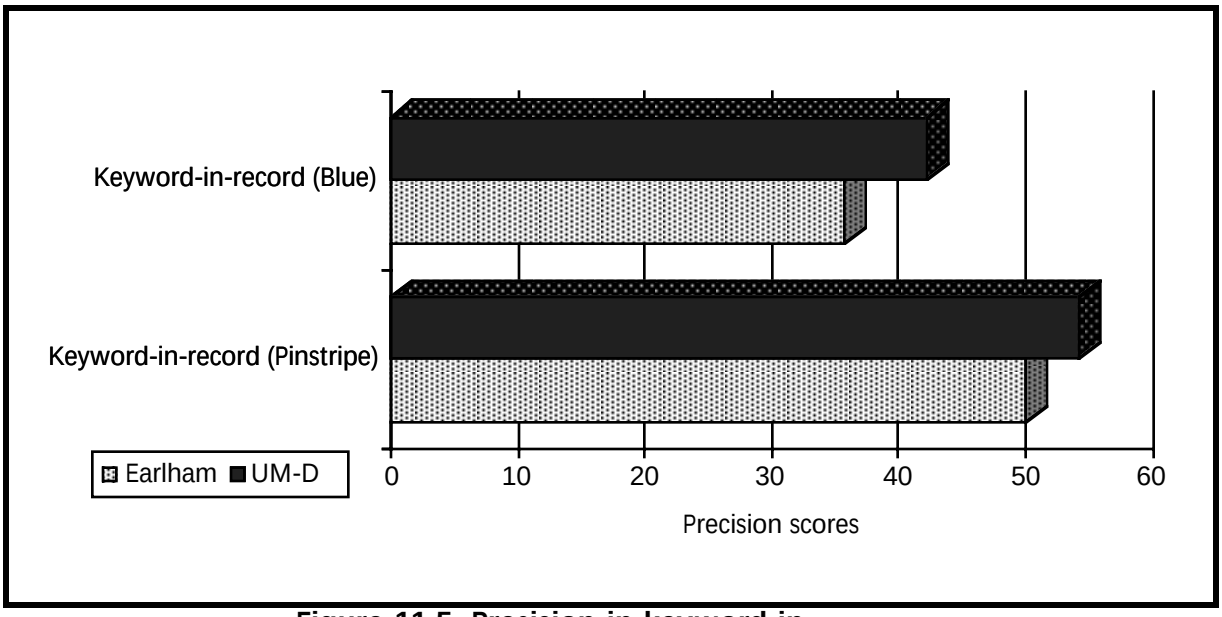

Figure 11.5. Precision in keyword-in-record searches

Precision scores for keyword-in-record searches were higher in the Pinstripe System than in the Blue System. However, precision scores in the Blue System were based on the results of about a half dozen searches.

11.6.3 Successful and Unsuccessful Keyword-in-record Searches

Table 11.19 summarizes the number of keyword-in-record searches and reasons for success or failure.

Table 11.19. Keyword-in-record Searches

Reasons for success/failure	UM-D		Earlham	
	Blue	Pinstripe	Blue	Pinstripe
Successful searches:				
Simple search	1 (16.7%)	93 (52.0%)	5 (100%)	19 (44.2%)
Unsuccessful searches:				
Perseverance	1	6	0	4
Too many retrievals	0	17	0	3
Database failure	1	40	0	7
Query specificity	0	6	0	3
Vocabulary	3	9	0	7
Spelling problems	0	2	0	0
Search approach failures	0	6	0	0
Subtotal	5 (83.3%)	86 (48.0%)	0 (0.0%)	24 (55.8%)
Total	6 (100%)	179 (100%)	5 (100%)	43 (100%)

Like keyword-in-subdivided-heading searches, keyword-in-record searches were not very common in the Blue System. This was due to the several searches preceding this search to which the search trees had submitted user queries and, probably, had been successful making matches. Of course, users could always expand their searches to retrieve the results of keyword-in-record searches. Keyword-in-record searches were much more common in the Pinstripe System. The keyword-in-record search was one of three subject searching approaches in this system.

About half of keyword-in-record searches were successful. Keyword-in-record searches were low posted in the Blue System where they averaged 2.3 titles per search. Keyword-in-record searches were much higher posted in the Pinstripe System where they averaged 81.8 titles per search. Searches ranged from zero to 2,324 titles. If postings from the twenty searches in which users terminated their searches without rating even the first title displayed were disregarded, keyword-in-record searches averaged 44.7 titles per search. This dramatic reduction in the average number of titles was due to several searches among the twenty searches that retrieved so many titles that users did not bother to display them.

On occasion, users rated several titles “very useful” that contained the same main or subdivided subject heading. A system that features relevance feedback could automatically examine subject headings in useful titles. Users who want additional material on their topics of interest could benefit from the system’s retrieval of titles

bearing these subject headings. Table 11.20 describes several keyword-in-record searches with such headings.

Table 11.20 Frequently-occurring Subject Headings in Keyword-in-record Searches

User query	Explanation
southern economy	Blue System retrieves 2 titles. User rates one title "very useful" bearing the subdivided subject heading "Southern states — Economic conditions."
slavery and black women	Blue System retrieves 2 titles. User rates one title "very useful" bearing the subdivided heading "Afro-American women — History — 19th century." Examples of titles retrieved with this subject heading are *Collected Black women's narratives* and *We are your sisters: Black women in the 19th century.*
black families	Pinstripe System retrieves 19 titles. User rates 9 titles "very useful;" all bear the subject heading "Afro-American families."
automatic transmission	Pinstripe System retrieves 7 titles. User rates 1 title "very useful" bearing the subdivided subject heading "Automobiles — Transmission devices."
code	Pinstripe System retrieves 16 titles. User rates 3 titles "very useful;" all bear the subject heading "Pulse-code modulation."

Perseverance was not as much of a problem in keyword-in-record searches as it was in controlled vocabulary searches. Eleven keyword-in-record searches (4.7%) were marred by perseverance. About twice as many (20) searches were marred by a different type of perseverance — too many retrievals. These searches averaged 411 titles; they ranged from 61 to 2,324 titles. Users did not rate the first title that the systems automatically displayed in these high-posted searches nor did they complete post-search questionnaires. They just walked away.

The failure of 48 (20.6%) keyword-in-record searches was attributed to database failure. Most of these searches retrieved a few titles but they were false drops. Examples are keyword-in-record searches for "propeller" that retrieved one title entitled *Measuring and metering of unsteady flows* containing the word "propeller" in a contents note, i.e., "Responses of small propeller current meters to impulsive and pulsating turbulent flows," and for "black church history" that retrieved one title entitled *Quest for equality: the life and writings of Mary Eliza Church Terrell, 1863–1954* bearing the words "black" and "women" in the series title (i.e., "Black women in United States history") and the word "church" in the title.

11.7 Subject Searches for Personal Names

11.7.1 Review of Personal-name Searching

Prior to allowing users to enter subject queries, the experimental online catalogs asked users whether their queries involved a person's name. The catalogs asked this question in the same way in the Blue and Pinstripe Systems (see figure 11.6). The catalogs responded to users who gave positive answers to this question with a series of pop-up screens that prompted users to enter surname, given name, and topic (see figures 5.29–5.31).

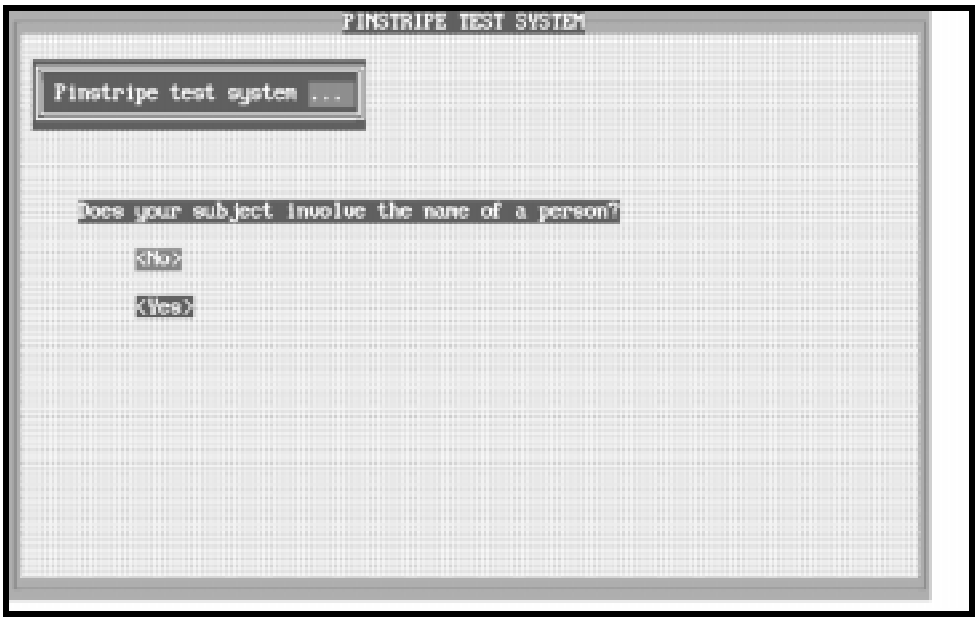

Figure 11.6. Asking users whether their queries involve names

End users entered subject searches for personal names into the experimental online catalogs at UM-Deaborn and Earlham College. The database for the former was composed of library materials in computer science and technology. Few personal names were assigned to UM-D bibliographic records. In fact, the entire alphabetical list contained less than 150 personal-name subject headings.

The database for Earlham College was composed of library materials in American history. Many names were assigned to bibliographic records. In fact, the alphabetical list of personal-name subject headings contained almost as many unique names as the alphabetical list of topical subject headings.

11.7.2 Queries for Personal Names

11.7.2.1 Unusable Queries at UM-D

Despite the infrequency with which personal names were connected with computer science and engineering topics in the UM-D database, some UM-D users responded positively to the system's question about personal names and entered queries bearing personal names. Examples were:

- limbrough [sic] (surname), rush (given name)
- pasteur (surname), louis (given name)
- regan [sic] (surname)
- corniliustacitus [sic] (surname)
- kevorkian (surname), physician assisted suicide (topic)
- perot (surname), ross (given name), petroleum (topic)
- jakeson [sic] (surname), michael (given name), singing muise [sic] (topic)
- ogata (surname), control theory (topic or title)

These eight queries were typical of the personal-name queries for subjects that users entered into the experimental online catalog at UM-D. Most queries were out of scope (first six queries listed above). A few queries were characterized by playing around (seventh-listed query above) or author-title searches (last-listed query above).

Two subject queries for personal names that UM-D users entered into the experimental online catalogs deserve mention, "pascal" and "gauss." The former is a name of a computer programming language or a unit of pressure. The latter is an electromagnetic unit of magnetic induction. Both queries were valid *subject* queries. The systems submitted both queries to subject searching approaches for personal names because users responded positively to the system question about personal names being involved with queries. In fact, both queries did involve names of persons, Blaise Pascal, a seventeenth-century French scientist, and Johann Karl Gauss, a nineteenth-century German mathematician.

The question about names in queries that users answered prior to entering their queries was quite effective in categorizing subject queries as personal names or subject queries generally. In view of the two queries discussed above (i.e., "pascal" and "gauss"), it was not entirely effective in effecting a correct categorization. Future computer systems that separate subject searches for personal names or for subject queries generally could

experiment with different wording for the question that categorizes subject queries as personal names or subject queries generally.

A handful of UM-D users answered negatively to the question about names in queries. They entered personal names (and, sometimes, title words) into the catalogs. Examples were:

- assad
- hitler
- bush, g
- sinclair upton
- schaum's
- schaum's outline series dynamics
- dijkstra

Some names were out of scope (first four names listed above). Other names were probably attempts at author or author-title searches in the areas of computer science and engineering (last three names listed above). Why did users answer negatively to the question about personal names that resulted in the system's submission of these names to subject searching capabilities for subjects generally? In the absence of post-search interview data, we can only speculate on such reasons. Some users could have changed their minds and entered different queries than the ones they originally intended to enter. Other users could have misunderstood the question about personal names. Still other users could have failed to read the question about personal names.

Future computer systems that separate subject searches for personal names or for subject queries should experiment different wording for the question that categorizes subject queries as personal names or subject queries generally. Users who become frequent users of a particular online catalog could become bored with answering the same question over and over again. They might need a different method of submitting queries for personal names than users who use the system infrequently.

11.7.2.2 Unusable Queries at Earlham

Personal-name subject headings were quite common in the bibliographic database on American history at Earlham. Yet findings about the entry of queries for personal names were quite similar to those at UM-D. End users at Earlham entered queries for personal names that were out of scope. Some could have been author-title searches for known items. Examples were:

- calvin (surname), john (given name), treatises against the anabaptists (topic)
- chopin (surname), kate (given name)
- poe (surname), edgar (given name)
- schlesinger (surname), arthur (given name), the age of jackson (topic or title)
- arendt (surname), hannah (given name), totalitarianism (topic)
- webster (surname), dictionary (topic or title)

On occasion, end users at Earlham answered negatively to the system's question about personal names in queries. Consequently, the system submitted their queries to subject searching capabilities for subjects generally. Below are examples of out of scope queries bearing personal names that were submitted to such capabilities:

- tanner, henry
- marilyn monroe
- rodin

11.7.2.3 Elements in Usable Queries at Earlham

Users entered 27 different queries for personal names into the experimental online catalogs at Earlham. Some queries contained topical elements:

- jackson (surname), andrew (first name), jacksonian democracy (topic)
- calhoun (surname), john (given name), state rights (topic)
- sumner (surname), charles (given name), reconstruction (topic)

Other queries only contained name element(s):

- roosevelt (surname), eleanor (given name)
- rustin (surname), bayard (given name)
- fremont (surname)

Of the 27 queries for personal names, a few deserve special mention because of ways in which users stated name and topic elements. Both Blue and Pinstripe Systems prompted users for surname, given name, and topic elements. Although they did not prompt users for middle name elements, four queries contained two elements in the given name pop-up box that represented first and middle name elements. These queries were:

- king (surname), martin luther (first and middle names entered into first name pop-up box), about the bycotts [sic] they had way back in mississippi (topic)
- king (surname), coretta scott (first and middle names entered into first name pop-up box), learn about what happened in 1927 (topic)
- king (surname), martin luther (first and middle names entered into first name pop-up box), about in the 1925 (topic)
- washington (surname), booker t (first name and middle initial entered into first name pop-up box), african american history (topic)

The first three queries were entered successively into the experimental online catalog. They were probably entered by the same individual. In view of our experience of the four queries above, we could speculate that some users may feel compelled to enter the middle name or initial for famous people that are known by their full names. Yet, in the following query for an author who is known by her full name (i.e., Harriet Beecher Stowe), the user did not enter the middle name:

- stowe (surname), harriet (first name), uncle tom's cabin (topic)

Two other queries were especially interesting:

- hoover herbert (surname and given name entered into surname pop-up box), herbert (first name), economic (topic)
- robert lafollette (given name and surname entered into surname pop-up box), mcarthy (topic)

In the first-listed query above, the user entered both surname and given name in inverted form into the surname pop-up box. He repeated the entry of the given name in the given name pop-up box. Perhaps he realized he entered both first and last names into the surname pop-up box after the given name pop-up box prompted him to enter the given name. When the user was ready to submit the entire query to the system, the system gave the user the opportunity to change entered name and topic elements but he did not make changes. The Pinstripe System performed a keyword-in-record search for name and topic elements in this query and was unsuccessful in retrieving titles. The Blue System was also unable to retrieve titles for name and topic elements but it continued searching by placing the user at the personal-name subject heading in an alphabetical list of personal-name subject headings beginning with the name "Hoover, Herbert, 1874–1964." Unfortunately, this Blue-system search was marred by user perseverance.

In the second query, the user gave the name in direct form. The Pinstripe System performed a keyword-in-record search for name and topic elements and was unsuccessful in retrieving titles. The Blue System was also unable to retrieve titles for name and topic elements but it continued searching by displaying to the user an alphabetical list of personal-name subject headings that began with the letter "r." The user browsed backward repeatedly and found the subject heading "LaFollette, Robert Marion, 1855–1925." She selected this subject heading, displayed eight titles, and gave two titles "very useful" ratings.

The ASTUTE experimental online catalog was not unique in its division of subject searches into subject searches for personal names and for subjects generally. UCLA's ORION online catalog also enlists such a division. The ASTUTE experimental online catalog was unique in its prompting of users for the name and topic elements of personal-name queries. Additional research is needed to determine how users enter a wide range of names, especially names in which the distinctions among surname, given-name, and middle-name elements are not as clear as the names entered by Earlham College users, for example, foreign names (Vincent Van Gogh, Jan van Eyck, Rogier van der Weyden) Native American names (Tecumseh, Blue Jacket, Geronimo), Classical names (Hippocrates, Aristotle), and to determine whether systems can provide satisfactory results.

11.7.2.4 Subject Searching Approaches for Personal-name Queries

When users entered subject queries for personal names, the Blue System was governed in its selection of a subject searching approach by the search trees for personal-name queries (see Table 4.1). The system first submitted queries with name and topic elements to the keyword-in-record search. If it failed to retrieve titles bearing these elements, it ignored the topic element and responded with the results of an alphabetical search in which listed subject headings for personal names were in the same alphabetical neighborhood as the last and first name elements of the query. The Pinstripe System always responded to personal-name queries with the keyword-in-record search.

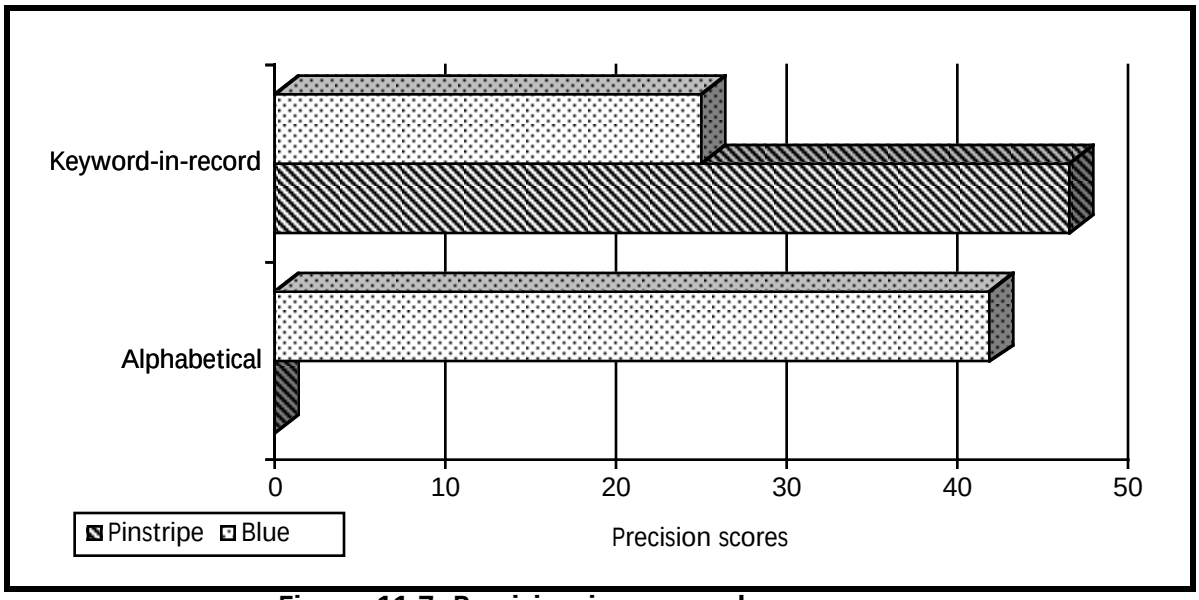

Figure 11.7. Precision in personal-name subject searches

Figure 11.7 shows precision scores for personal-name subject searches. (Precision scores include conflicting relevance assessments and relevance assessments for repeated displays of the same titles.)

Precision scores for keyword-in-record searches were higher in the Pinstripe System than in the Blue System. Precision scores for the Blue System's alphabetical approach and the Pinstripe System's keyword-in-record search were about the same.

Table 11.21 summarizes the subject searching capabilities that handled subject queries for personal names in the Blue and Pinstripe Systems and characterizes them as successful or unsuccessful.

Table 11.21. Subject Searches for Personal Names

Reasons for success/failure	Earlham	
	Blue	**Pinstripe**
Successful searches:		
Keyword-in-record	4	8
Alphabetical	8	n/a
Subtotal	12 (52.2%)	8 (38.1%)
Unsuccessful searches:		
Perseverance	6	0
Database failure	4	4
Search approach failures	0	9
Query specificity	1	0
Subtotal	11 (47.8%)	13 (61.9%)
Total	23 (100%)	21 (100%)

Over half (12 of 23) of subject searches for personal names were successful in Earlham's Blue System. About a third (8 of 21) of subject searches for personal names were successful in Earlham's Pinstripe System. The most frequent reasons for failure in the Blue and Pinstripe Systems were perseverance and search approach failures, respectively. Both systems had their share of database failures.

Of the 27 subject queries for personal names, all but four queries featured elements for names and topics. The keyword-in-record search failed to retrieve titles for half of these queries because of the topic element. Frequently, this resulted in a search approach failure for the Pinstripe System (see section 10.3.1.6). For example, a user entered the query "sumner" (surname), "charles" (first name), "reconstruction" (topic) into the experimental online catalog. The Pinstripe System submitted this query to a keyword-in-record search and failed to retrieve titles. The Blue System also submitted this query to a keyword-in-record search. When it failed to retrieve titles, the system ignored the topic element and responded with the results of the alphabetical approach. The user selected the personal-name subject heading "Sumner, Charles, 1811–1874" from the alphabetical list, retrieved eight titles, and rated all eight titles "very useful."

Search approach failures were quite common for the personal-name queries bearing topic elements that users entered into the Pinstripe System. The failure of nine subject queries was attributed search approach failures in the Pinstripe System.

Several Blue-system searches were successful in retrieving titles through the alphabetical approach to which users gave positive relevance assessments even though the titles did not specifically refer to the topic element of their queries. The failure of a few Blue-system searches to retrieve useful titles was attributed to user perseverance, query specificity, or database failure. For example, after the query "whittier" (surname), "john" (given name), and "abolition" (topic) failed to retrieve titles in a keyword-in-record search, the Blue System responded with the alphabetical approach but the alphabetical list of personal-name subject headings did not contain a subject heading for John Whittier. The failure of this Blue-system search was attributed to database failure.

On occasion, users were not satisfied with the results of Blue-system searches in which the alphabetical approach ignored the topical element of their queries and retrieved titles based only on the name elements of their queries. For example, in response to the query "king" (surname), "coretta scott" (first name), "learn about what happened in 1927" (topic), the Blue System ignored the topical element because of its inability to retrieve titles through a keyword-in-record search and responded with an alphabetical list of personal-name subject headings. The user selected the personal-name heading for Coretta Scott King, retrieved several titles, and gave them all negative relevance assessments. Perhaps this user would have only been satisfied with titles that specifically referred to Mrs. King in 1927.

11.7.2.5 Recommendations for Personal-name Queries

The Blue System performed much better than the Pinstripe System in terms of responding with a subject searching approach that led users to titles on their topic of interest or on a facet of their interests. Search approach failure occurred in the Blue System with much less frequency; however, the failure of a few Blue-system searches was attributed to user perseverance, query specificity, or database failure.

We recommend the implementation of the search trees for subject queries for personal names in operational online catalogs. The search trees should be accompanied by pop-up windows that prompt users for the elements of their personal-name queries. This prompting enables systems to differentiate between name and topic elements of queries and choose to ignore certain elements in their attempts to retrieve titles. More research is needed to ensure that users enter the correct elements in various pop-up windows but this should not deter system designers from implementing the search trees and prompting mechanisms in operational systems.

The procedures for prompting users for personal names in subject searches should be extended to prompting users for personal names in author searches and for personal

names and titles in a combination author-title search. This could help eliminate the frequency with which users enter direct forms of names in author searches into systems that expect indirect forms of names in author searches and respond with an alphabetical list of author names based on the left-most name element in user queries. It could also help systems make informed decisions about the particular element(s) they should ignore in their attempts to retrieve titles. For example, if systems fail to retrieve titles using both author name and title elements of author-title queries, they could ignore the author or title elements of queries in their attempts to retrieve titles.

11.8 Librarians' Subject Searching Experiences

11.8.1 Introduction

Eight librarians participated in the Comparison Search Experiment. One librarian was a cataloger and seven librarians were reference librarians. They accounted for a total of thirty-three search administrations. The analysis of subject searches in chapters 7–11 intermingled searches performed by end users and searches performed by librarians. We did not separate results for librarians from results for end users generally.

Comparison Search Experiments with librarians included personal interviews following their searches in the Blue and Pinstripe Systems. We placed their remarks into the following broad categories: (a) subject searching capabilities, (b) number of retrievals, (c) expand capability, (d) browsing, (e) search trees, (f) search administration problems, (g) database, (h) personal-name searching, and (i) housekeeping features. Most comments were paraphrased versions of the comments librarians made. Comments placed in quotation marks were actual statements made by librarians.

11.8.2 Categorized Comments

a. Subject searching capabilities

a1. General comments

- I am waiting for the system to tell me what types of subject searches I can do. I like search menus like Mirlyn has with searches for s= (subject heading browse search) and k= (keyword search with implicit Boolean).
- When the system finds only a few titles for your query, it should have capabilities for finding more titles. (2 librarians)

- System should have been able to ignore the "action" word in my query, "origins of the ku klux klan," and retrieved titles for the "ku klux klan" component of the query.
- Following a keyword-in-record search with few retrievals, the system should have suggested more subject searching terms to help me find more titles. (2 librarians)

a2. Boolean searching

- It is important for me to know if I can search for more than one term at a time or if I can combine terms.
- The system could not handle a Boolean query, i.e., ((“blacks or afroamericans or chinese) and west).” Add a Boolean searching capability because some of our students know how to construct Boolean search statements.
- I am unsure whether the system will allow explicit truncation.

b. Number of retrievals

- Our library's online catalog shows you individual postings for your query words so you have an idea which word is detrimental to your query.
- Add the number of retrievals for listed subject headings so I know whether I want to limit my search.

c. Expand capability

- I like the expand function because it helped me find more titles on the topic I wanted. (4 librarians)
- I expect that the "expand" capability finds more titles using broader and narrower terms. (3 librarians)
- By choosing "expand," I had hoped to retrieve a list of broader and narrower terms. Instead, the system responded with the same titles that I saw previously.
- I like the expand function. It didn't help in my search today but I might need it in the future. (2 librarians)
- "Expand" capability seems like a "black box." I sure would like to know what it is doing.

d. Browsing

- I liked the keyword-in-subdivided-heading list because listed terms allowed me to narrow my topic or find a more precise term for my topic of interest.

- I liked the option to browse broader and narrower terms.

d1. Alphabetical lists

- When my search fails to retrieve titles, the catalog should show me a list of subject headings to help me find relevant searching terminology. (4 librarians)
- I liked when the system responded with an alphabetical list of subject headings; however, I was perplexed when it didn't respond with this list. I think searchers would find this inconsistency perplexing also.
- I preferred the Pinstripe System because it responded to my query "human powered vehicles" and "virtual reality" with an alphabetical list of terms. I found the terms "Human locomotion" and "Virtual computer systems" which are somewhat close to my interests. The Blue System merely told me it could not find titles on my topic of interest. (2 librarians)
- Neither system found titles for my query "computer-assisted instruction. " At least they could have responded with an alphabetical list of terms.
- I like to look at an alphabetical list of subject headings to find more titles. I wish this system had such a capability. (2 librarians)

d2. Broad categories

- Blue System makes me go through too many levels of selecting subdivided terms before I am able to display titles. In my search for "civil rights," I had to select four or five levels of terms before I could display titles assigned to the subject headings "Civil rights — United States — History — 20th century" and "Civil rights demonstrations — Florida — St. Augustine — History — 20th century." And then I only retrieved one or two titles! If I'm only going to retrieve two titles, can't I display these titles sooner?
- I would like to see subdivided and unsubdivided subject headings in a single list. The other approach (broad categories) is too complicated for selecting subdivided subject headings.
- If searchers do not know LCSH, the broad categories list could help them.

e. Search trees

- "As a librarian, I want to know what the system is doing or I want to tell it what to do." The system should give me a clue as to what search to do next to produce retrievals.
- It bothers me that I don't know what fields the system is searching. I will sometimes do a title search and scan subject headings in retrieved titles to find a correct subject heading but I cannot perform this strategy because I

do not know how the system searches. “I like to be in control when I’m searching.”

- I would like the system to give me a message telling me what it is doing. I don’t know if this would be as much help to students as to librarians but students do pick up on what systems are doing.

f. Search administration problems

- I wanted to display titles but I must have hit the wrong key because the system switched systems on me.

g. Database

- I like the idea of having a system dedicated to computer science and technology because it is easy to block out extraneous material.

h. Personal-name searching

- In my search for “willke” (surname), “wendell” (given name), ”nomination” (topic), the system retrieved general titles on Wendell Willke but I would have preferred titles specifically on his nomination.
- In my search for “jefferson” (surname), “thomas” (given name), ”inventions” (topic), the system retrieved general titles on Thomas Jefferson but I would have preferred titles specifically on his inventions.

i. Housekeeping features

- I like the system’s displays of retrieved titles (2 librarians)
- I like how the system blinks “Searching ...” as long as it is still searching for your terms.
- I like having both keyboard and mouse so I can choose between them.
- I like how the system gives me a chance to edit my query before I enter it.
- I like how the system suggests that my query might be misspelled. (2 librarians)
- I am not familiar with how the system indicates disabled options. Perhaps it would be better to limit displayed options to those that I can select.
- I don’t like how you have to hit the <Enter> key twice to enter or select a search term.
- I like the windowing capabilities.
- I like how few keystrokes I have to enter to perform a search.

11.8.3 Suggested System Improvements

Librarians' comments call for the following improvements to systems that feature search trees:

- A capability for "set detail on." Such a capability would be available to librarians and end users who want to know how the system is processing their queries. This capability could reduce the mystery of both the search trees and expand capability because both submit user queries to various subject searches but do not report unsuccessful results.
- User feedback capability. User feedback could enlist several approaches to retrieving additional titles, especially following low-posted keyword searches. For example, systems could perform searches for subject headings common to the many titles that users retrieve and give "very useful" ratings.
- Best-match approach. The best-match approach would feature stemming, weighted-term probabilistic retrieval, and output ranking. Systems would perform searches on a best match, combinatorial basis. A match on fewer than all query-word stems may retrieve titles but titles that have all the query-word stems would be displayed first. This approach would have helped librarians retrieve titles for the query "origins of ku klux klan" query and Boolean query "((blacks or afroamericans or chinese) and west)." It would also address comments on finding additional titles following low-posted keyword searches.
- Rename the "expand" capability. When librarians became familiar with the operation of the "expand" capability, they found it valuable for retrieving additional titles on occasion. Some remarks, however, indicated that librarians thought this capability would retrieve lists of broader and narrower terms. This capability should be renamed to make it indicative of its function.
- Indicate titles retrieved and displayed previously. Blue- and Pinstripe-system users had several opportunities to retrieve and display the same titles repeatedly in subject searches. Both systems should indicate titles that users have retrieved in previous searches and displayed in previous searches. In a system that features icons, a check mark or an open eye could be placed adjacent to listed titles to indicate titles that users have retrieved and displayed in previous searches, respectively.
- Enlist the alphabetical approach as the search "of last resort." When the search trees fail to retrieve titles, systems should respond with the

alphabetical list of subject headings to give users ideas of terms in the alphabetical neighborhood of their entered term.

- Reduce the complexity of exact-approach displays. If exact-approach searches result in the retrieval of a few subdivided subject headings, systems should display these headings immediately instead of requiring users to proceed through several levels of broad categories to display subject headings and, ultimately, to retrieve titles.

Greater experience with the Blue System would make several aspects of the experimental system less uncommon or more familiar: (1) option disabling, (2) expanding, and (3) responding to pop-up menus to differentiate between more than one action following the entry or queries or selection of terms from lists.

11.9 Chapter Summary

Chapter 11 examined subject searching capabilities in the Blue and Pinstripe Systems. It focused primarily on the success of these capabilities but it also covered frequent reasons for their failure.

Controlled vocabulary searches were especially plentiful in the Blue System and accounted for almost two-thirds of searches conducted in the Blue System (section 11.2). About half (52.5%) of controlled vocabulary searches were successful in retrieving at least one title rated "very useful" by system users. Many (54.5%) of these successful searches were searches characterized by no browsing (see Table 11.5). That is, users displayed retrieved titles and terminated their searches without browsing available options for broader or narrower terms, subdivided terms, notes, or expanding the search. Only 14.4% of successful controlled vocabulary searches featured extensive browsing.

Frequent reasons for the failure of controlled vocabulary searches in the Blue System were perseverance and navigation problems (see sections 11.2.4 and 11.2.5). Some of the searches in which perseverance was a problem could have been navigation problems. Based on the analysis of searches that failed due to perseverance and navigation problems, the following suggestions were made for improving the exact search:

- Reducing the complexity of the exact-approach main menu.
- Consolidating several exact-approach main menu options into fewer options based on numbers of retrievals.

- Including end users in the design and pretesting of the exact-approach main menu.
- Pretesting the wording of exact-approach main menu options with end users.
- Designing and implementing an easy, understandable backtracking option.

The keyword-in-subdivided-heading search was a controlled vocabulary search but it was treated separately from other controlled vocabulary searches because it was independent of the exact approach (section 11.3). Keyword-in-subdivided-heading searches were much more frequent in the Pinstripe System than in the Blue System. The reason why only nine of 156 keyword-in-subdivided-heading searches were conducted in the Blue System was attributed to the search trees that submitted queries to the keyword-in-subdivided-heading search after submitting queries to three other controlled vocabulary searches (section 11.3.1).

Two-thirds (29 of 44) of successful keyword-in-subdivided-heading searches featured no browsing (see section 11.3.3). These were simple searches in which users selected one listed subject heading, displayed one or more titles, and gave "very useful" ratings to one or more titles. Only one of 44 successful keyword-in-subdivided-heading searches featured more than a little browsing.

Unsuccessful keyword-in-subdivided-heading searches outnumbered successful the keyword-in-subdivided-heading searches (see section 11.3.4). Almost three-quarters (112 of 156) of keyword-in-subdivided-heading searches were unsuccessful. Frequent reasons for failure were perseverance, database failure, and searches that retrieved no subject headings.

Title-keyword searches were unique to the Blue System (see section 11.4). Successful title-keyword searches outnumbered unsuccessful title-keyword searches. Almost 70% of title-keyword searches in UM-D's Blue System were successful; over 85% of title-keyword searches in Earlham's Blue System were successful. On occasion, users rated several titles "very useful" that contained the same main or subdivided subject heading. A system that features relevance feedback could automatically examine subject headings in useful titles. When title-keyword searches retrieve few titles, users could find additional material on their topics of interest from the system's retrieval of titles bearing these subject headings.

Both Blue and Pinstripe Systems featured keyword-in-record searches but the majority (95.3%) of keyword-in-record searches were conducted in the Pinstripe System (section 11.6). About half of keyword-in-record searches were successful (see Table 11.19). Keyword-in-record searches were low posted in the Blue System where they averaged

2.3 titles per search. Keyword-in-record searches were much higher posted in the Pinstripe System where they averaged 81.8 titles per search. Large retrievals were the major problem with keyword-in-record searches in the Pinstripe System.

No subject queries for personal names that users entered into the experimental online catalog at UM-D were valid names (section 11.7). Most were out-of-scope or playing around. Users also entered subject queries for personal names that were deemed invalid into the catalog at Earlham; however, a total of 27 valid queries were entered. All but four of these queries contained topical elements. Blue-system search trees for handling personal-name queries were especially effective.

Post-search interview with librarians resulted in several suggestions for improving search trees. The most important suggestions were: (1) add a user feedback capability based on titles users rate “very useful,” (2) add the best-match approach, (3) reduce the complexity of exact-approach displays, and (4) enlist the alphabetical approach as the search “of last resort.”

The Blue System performed much better than the Pinstripe System in terms of responding with a subject searching approach that led users to titles on their topic of interest or on a facet of their interests. Search approach failure was typical of Pinstripe-system searches for personal names, especially when queries contained topic elements. The failure of a few Blue-system searches for personal names was attributed to user perseverance, query specificity, or database failure.

12 Redesigning Search Trees for Responsive Subject Searching

12.1 Introduction

Chapter 12 focuses on searching problems that have significant impact on the redesign of search trees — spelling and vocabulary problems. The bulk of the chapter is given to reconfiguring the search trees based on findings from the online retrieval tests that previous chapters feature (chapters 7–11).

12.2 The Need for Automatic Spelling Correction

The ASTUTE experimental online catalog did not feature automatic spelling correction. It did, however, check user queries for possible spelling errors, inform users of query elements that failed to produce retrievals, and suggest users check such elements for spelling errors.

The Blue System checked user queries for subjects generally following the system's attempt to make a match in exact and alphabetical approaches. It compared words in user queries with words in a keyword index to determine whether the individual words in queries were used in the database. It checked the words in queries from left to right. An example is the user query "noegro peospirity in late 1920's." The system failed to find the word "noegro," informed the user of its failure, and suggested the user check spelling (see figure 12.1). The user corrected this word and the new query he submitted to the system was "negro peospirity in late 1920's." The system failed to find the word "peospirity," informed the user of its failure, and suggested the user check spelling (see figure 12.2). After responding to several system prompts to check spelling, the user eventually, the entered a query that contained no spelling errors, viz. "negro prosperity in late 1920's."

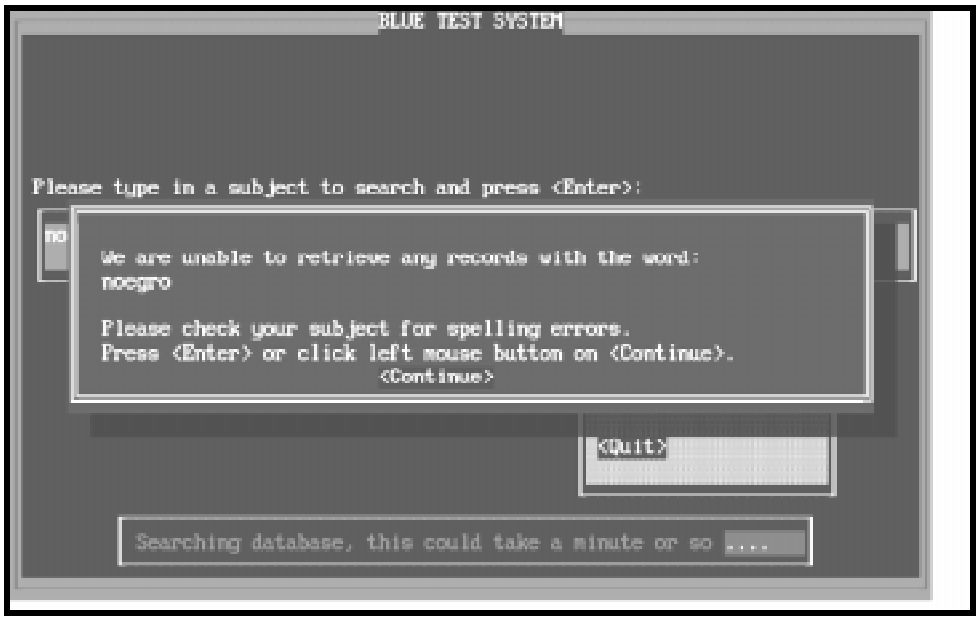

Figure 12.1 Informing users of possible misspellings

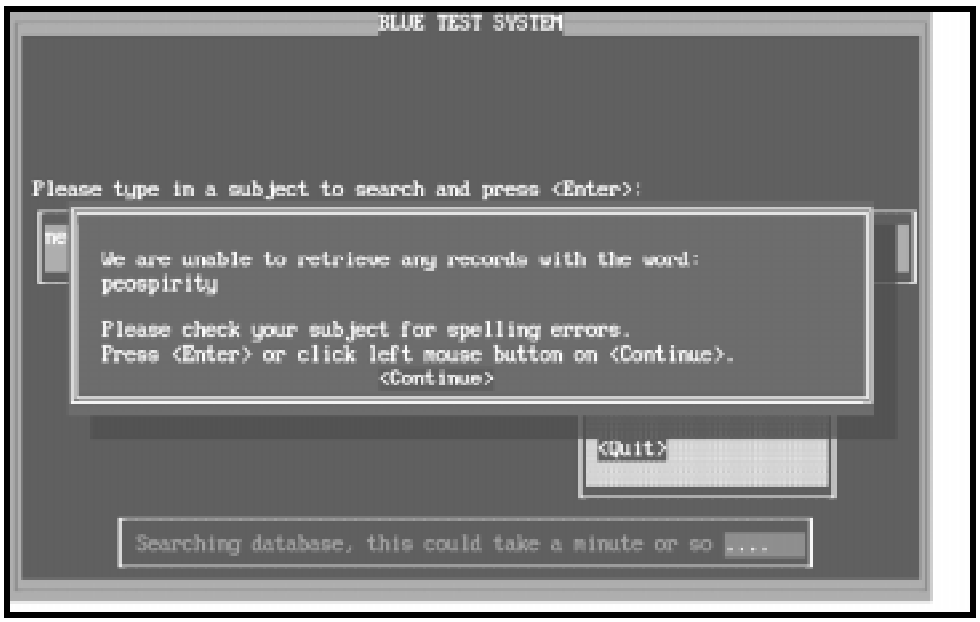

Figure 12.2. Checking queries from left to right for possible misspellings

When the Pinstripe System's random selection algorithm called for the alphabetical approach, the system made no attempt to find possible spelling errors. When it called for keyword-in-subdivided-heading or keyword-in-record searches, the system performed the same error-checking routine as the Blue System. That is, it checked a

keyword index to determine whether the individual words in queries were used in the database. It checked the words in queries from left to right, informed users of query elements that failed to produce retrievals, and suggested users check such elements for spelling errors.

The experimental online catalogs responded to 134 queries with the message in figures 12.1 and 12.2 that informed users that their queries contained possible spelling errors. Table 12.1 describes what users did next.

Table 12.1. User Actions Following System Message Regarding Possible Spelling Errors

	UM-D		Earlham	
User Actions	**No.**	**%**	**No.**	**%**
Enter query on different topic	34	27.2	1	11.1
Enter same query	27	21.6	0	0.0
Quit search	15	12.0	2	22.2
Correct spelling	11	8.8	5	55.6
Enter same query minus unposted word	12	9.6	0	0.0
Enter same query and add new word	11	8.8	0	0.0
Enter new query with same stem as previous query	7	5.6	0	0.0
Enter singular or plural form of previous query	7	5.6	0	0.0
Enter acronym or spell it out	1	0.8	1	11.1
Total	125	100.0	9	100.0

Large numbers of users entered queries on different topics. Examples of queries bearing unposted words and the queries users entered following the system's message informing them of a possible spelling error are listed in Table 12.2. Words in italics were the unposted words that the experimental online catalog displayed to users for their correction.

Table 12.2. Succeeding Queries on Different Topics

Possibly Misspelled Queries	Next Queries on Different Topic
internet	usenet
microcad *graphing*	database
radiator design	heat transfer
zirconia	chemistry
general *relativity*	computers
The *z8* microcomputer	zilog microcomputers

Following the system's message informing them of a possible spelling error, large numbers of users entered the same query one or more times. Examples of such queries are "internet," "hovercraft," "androids," "barcode," "reinforce concrete," and "nanotechnology." Perhaps such users reentered the same queries because they wanted to make absolutely sure that the system had no titles on these topics. Reentering such queries, users might have been saying to themselves, "There's got to be information on this topic in here somewhere."

Following the system's message informing them of a possible spelling error, fifteen users at UM-D and two users at Earlham quit searching.

A total of sixteen users corrected the misspellings in their queries. Table 12.3 gives examples of misspelled queries and their corrected forms. Words in italics were unposted words that the experimental online catalogs suggested to users were misspelled.

Table 12.3. Misspelled Queries

Possibly Misspelled Queries	Corrected Queries
human power vechicles	human power vehicles
communcations	communications
carsuspension and handling	car suspension and handling
chemistryy	chemistry
abolishionism	abolitionism
c programming lanuage	c programming language
monitring performance of telephone operators	monitoring performance of telephone operators

Of the total 134 queries in which the experimental online catalog detected unposted words, 28 queries actually contained misspelled words. Users corrected sixteen of these queries. Examples of queries that users did not correct were "elecctronics," "circuts," "assemlby language," "cobal languages," and "ei\\EIFFEL." In response to the system's message informing them of a possible misspelling, some users quit searching, other users

entered different queries, and still other users added new words or deleted the possibly misspelled words from queries.

A total of 23 users responded to the system message about possible misspellings by entering the same query minus the unposted word or by adding a new word to the same query. Table 12.4 lists user queries that prompted ASTUTE to respond with the possible-misspellings message and subsequent user queries with deleted or added words. Words in italics were unposted words that the experimental online catalogs suggested to users were misspelled.

Table 12.4. Queries with Words Added or Deleted

Original queries	Added or deleted words in subsequent queries
rotation of axis	rotation of axis formula
probability and *statisics*	probability
barcode	2d barcode
equilizer filters	filters
graphing	microcad graphing
cache memory	memory
internet	internet network
three-dimentional dynamics	dynamics

A handful of queries were placed in remaining categories. Examples of succeeding queries that had the same stem as preceding queries were “automanual” and “auto,” “encoding” and “encode,” “architectual design” and “architecture.” Examples of succeeding queries that were singular or plural forms of preceding queries were “florida keys” and “florida key,” “air bag” and “air bags,” and “rotation of axis formula” and “rotation of axes formula.” Two queries contained acronyms, “IWW” and “International Workers of the World” and “cobal languages” and “common business oriented language.”

Frequently, the experimental system informed end users that their entered terms might be misspelled when, in fact, their entered terms were not posted in the database.

12.3 Vocabulary Problems and Search Trees

12.3.1 Introduction

The experimental online catalog contained information on the topics that interested several users. However, users phrased their queries in certain ways that prevented the

experimental online catalogs from making retrievals in response to their queries. When users entered queries with such phrasing, the two experimental systems responded in different ways. The Blue System exhausted the subject searching approaches on the search trees before responding to users that it was unable to find titles to satisfy their queries. The Pinstripe System performed the search using one of the three subject-searching approaches at its disposal before responding to users that it was unable to find titles to satisfy their queries. However, differences between the two systems in terms of subject-searching approaches were transparent to end users.

Chapter 10 (sections 10.2.1.6, 10.2.2.6, 10.3.1.5, and 10.3.2.6) discussed vocabulary problems from the perspective of failed searches in full or partial administrations of the search experiment at UM-D and Earlham. While this section features some of the search queries discussed in chapter 10, it focuses on how the search trees could be enhanced to handle queries with vocabulary problems.

12.3.2 Punctuation in Queries

A few queries failed to retrieve useful titles on users' topics of interest because of punctuation in user queries and/or controlled vocabulary terms. For example, the user query "computer aided instruction" failed to match the hyphenated subject heading "Computer-aided instruction" because we did not double post the hyphenated word in this subject heading as two words, i.e., "computer" and "aided," and one word, i.e., "computeraided."

The original specifications for the experimental online catalog called for double posting of hyphenated words (Drabenstott and Vizine-Goetz 1994, 242). Our failure to double post such words was an oversight in the design of the experimental online catalogs. This failure, however, brought to light problems with punctuation, especially hyphenated words. For example, if a term in a subject heading or subdivision was not hyphenated, the system would not generate two normalized forms of the subject heading, a form for the one-word variant and a form for the two-word variant, because the term contained no hyphen. On occasion, users might enter the term with a hyphen. Since the system did not double post the term, it would fail to make a match. For example, systems would normally generate only one form of the subject heading "Coffee growers," that is, the two-word normalized form "coffee growers." This subject heading has no *see* reference under "Coffee-growers," so systems would not double post this heading under "coffee growers" and "coffeegrowers." If a user entered the query "coffee-growers," systems would fail to make a match because it generated no index entry under the normalized form of this query, viz. "coffee growers."

Systems could anticipate this unfortunate situation by finding additional hyphenated forms of words and phrases in various fields of bibliographic records, e.g., title and series-title fields, chapter titles in contents fields. For example, if a title contained the hyphenated term “coffee-growers,” systems could search for subject headings or subdivisions bearing these two adjacent words and double post one-word and two-word normalized forms of these subject headings in the subject heading index, i.e., “coffee growers” and “coffeegrowers.”

Retrieval was adversely affected by our failure to omit punctuation from one other user query. The user query “c.a.s.e.” contained four periods. It retrieved several titles in a keyword-in-record search but all titles contained long contents notes with the initials “c., ” “a.,” “s.,” and “e.” If the periods and intervening spaces were deleted, the Blue System would match the *see* reference “CASE” for “Computer-aided software design” which probably described what the user had in mind.

12.3.3 Stemming User Queries

A case could be made for the inclusion of automatic stemming in the experimental online catalogs. Examples of queries that would have benefited from automatic stemming are described in Table 12.5.

Table 12.5. Stemming Queries to Enhance Retrieval

User query	Explanation
mechanical statics	Blue System exhausts subject searching approaches on the search trees and responds to the user that it is unable to find titles to satisfy the query. Pinstripe System submits this query to a keyword-in-subdivided-heading search and makes no retrievals. If the systems automatically stem the word "mechanical," they retrieve titles like *Engineering mechanics: statics, Schaum's outline of theory and probabilities of engineering: mechanics, statics, and dynamics,* and *Engineering mechanics: statics and dynamics.*
serial communication	Blue System exhausts the search trees without retrieving titles or subject headings. Pinstripe System's keyword-in-record approach fails to produce retrievals. If automatic stemming is applied to this query, the experimental system retrieves *Mastering serial communications* in a title-keyword search.
reinforce concrete	Blue System exhausts search trees without retrieving titles or subject headings. If this query is stemmed, the system retrieves several titles bearing the subject heading "Reinforced concrete."
bar coding	Blue System exhausts search trees without retrieving titles or subject headings. The alphabetical approach in the Pinstripe System places users in the alphabetical neighborhood of the subject heading "Bar codes." Also, stemming produces retrievals under this subject heading.

12.3.4 Best-match Approach

The best-match approach would be fairly successful in producing retrievals for user queries. Table 12.6 describes queries for which the best-match approach could yield retrievals. The best-match approach should also feature stemming to facilitate retrieval of variant forms of user query words.

Best-match algorithms typically give greater weight to low-posted words than to high-posted words. A new weighting system ought to be devised for library cataloging databases because subject content is represented in so few fields of bibliographic records, i.e., title and subject heading fields. This system should increase the weight given to frequently-occurring words in subject headings and subdivisions because matches on such words are likely to retrieve subject headings describing users' topics of interest or a facet of their topics of interest. For example, the stems "women" and "slave" occur more frequently in subject headings than the stems "their," "treatment," and "society." Table 12.6 lists three titles on the user's topic of interest that the former high-posted stems retrieve.

Table 12.6. Queries Suited to the Best-match Approach

User query	Explanation
flow through porous media	Blue System exhausts the search trees without retrieving titles or subject headings. Pinstripe System's keyword-in-record approach fails to produce retrievals. If this query is reduced to "porous media," the experimental system retrieves *Fundamentals of transport phenomena in porous media* in a title-keyword search.
fundamental of data structures in c	Pinstripe System's keyword-in-subdivided-heading approach fails to produce retrievals. If this query is reduced to "data structures in c," the experimental system retrieves *Data structures: an advanced approach using C* in a title-keyword search.
tolerance and geometric inspection	Both systems fail to produce retrievals. If this query is reduced to the stems of "tolerance" and "geometric," several titles are retrieved including *Geometric distancing and tolerancing* and *Tolerancing: computation on geometric uncertainties.*
black civil rights during the reconstruction	Both systems fail to produce retrievals. If this query is reduced to the stems of "black," "civil," "rights," and "reconstruction," titles such as *The Civil War and Reconstruction: a history of Black people in America, 1830–1880* and *The White response to Black emancipation: second-class citizenship in the United States since Reconstruction* are retrieved.
women slaves and their treatment by society	Both systems fail to produce retrievals. The stems of "women," and "slaves" retrieve titles such as *Women and the family in a slave society, Incidents in the life of a slave girl,* and *Six women's slave narratives.*
hopi indian land dispute	Both systems fail to produce retrievals. The stems of "hopi," "land," and "dispute" retrieve titles such as *The wind won't know me: a history of the Navajo-Hopi land dispute* and *The second long walk: the Navajo-Hopi land dispute.*

12.3.5 Go/see List

Another subject searching improvement would be the inclusion of a "go/see list" in online catalogs. Walker and Jones (1987, 68) described terms on the experimental Okapi system's "go/see list" as terms which the system treated as having the same meaning at search time. Examples of such terms were noun/adjective pairs, e.g., "French" and "France," acronyms and their spelled-out forms, e.g., "tv" and "television," and common, irregular plurals, e.g., "woman" and "women." When Okapi parsed user queries to find stopwords, it also searched for terms on the go/see list. Terms were not limited to one-word terms but sometimes included several words, e.g., "third world," "lesser developed countries," and "developing countries."

The research that led to the development of the search trees suggested a go/see list for one-word/two-word variations of concepts and noun/adjective pairs (Drabenstott and Vizine-Goetz 1994, 177, 292); however, ASTUTE was not equipped with go/see list functionality. In this research project, several queries could have benefited from a go/see list bearing alternative terminology, e.g., "blacks," "afro-americans," and "african-americans," "automobiles" and "cars."

12.4 Redesigned Search Trees

The analysis of online retrieval test data demonstrated that the search trees were more effective in selecting a subject searching approach that would produce useful information for the subjects users seek than users would select on their own. The search trees required enhancements to enable them to respond with useful retrievals to especially difficult user queries, that is, those for which the existing set of search trees produced no retrievals. This section presents the enhanced search trees.

12.4.1 Initial Search Tree

The search trees tested in this research project were described in chapter 4. The initial search tree remains basically unchanged from the original initial search tree (see figure 4.1). The only change from the original design is returning users whose searches failed to produce retrievals to the question about personal names. Based on user responses to this question, the new design of the initial search tree (figure 12.3) lets users distinguish their new or revised queries for personal subjects from queries for topical subjects generally and, based on summary characteristics that systems determine about the latter types of user queries, dispatches them to a particular search tree that favors certain subject searching approaches over others.

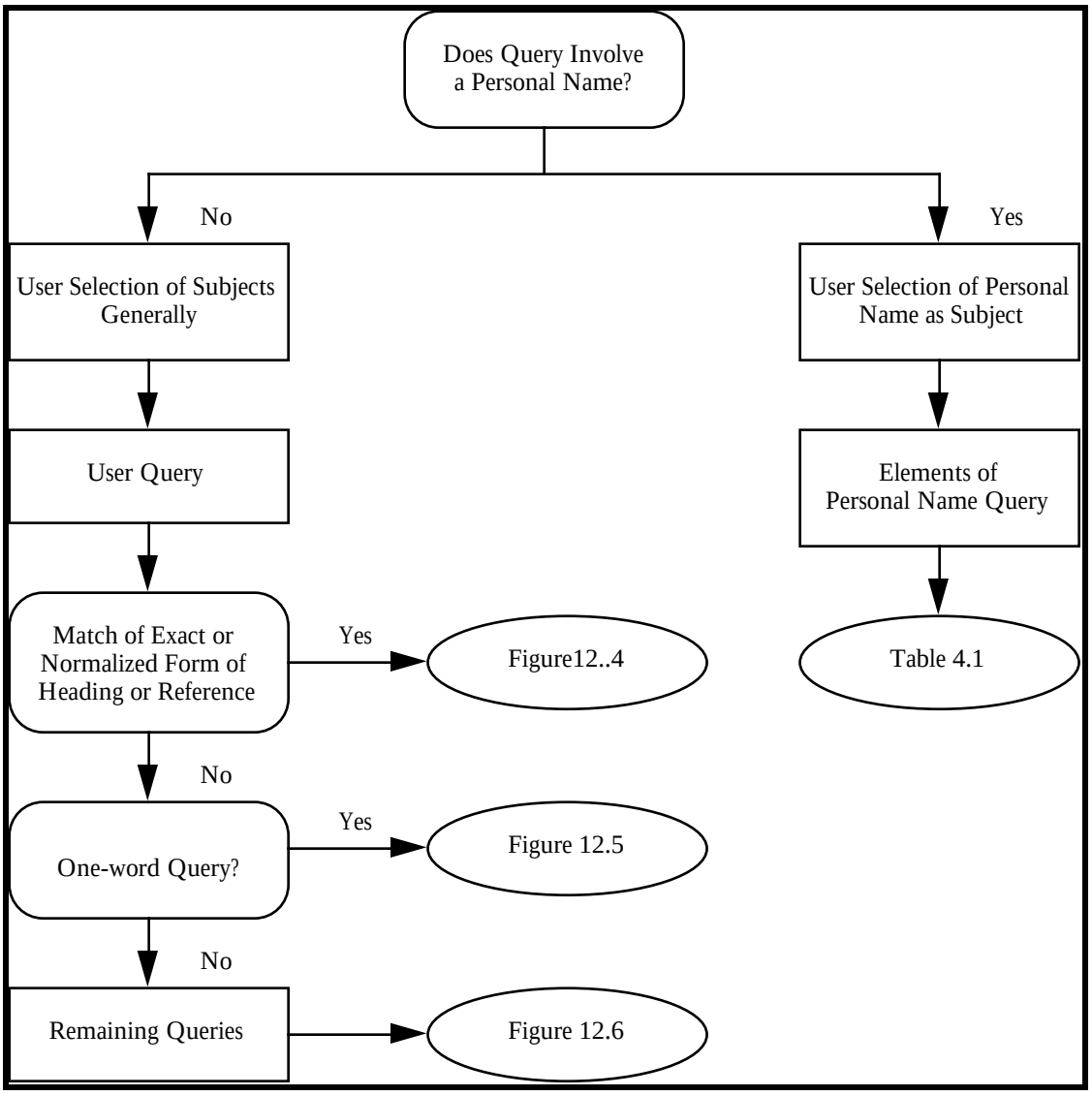

Fig. 12.3. Redesigned initial search tree

The search tree for the exact approach also remains unchanged. Systems respond to user queries for subjects generally that match exact or normalized forms of controlled vocabulary terms with the exact approach. Figures 12.4A and 12.4B show the search tree that features the exact approach.

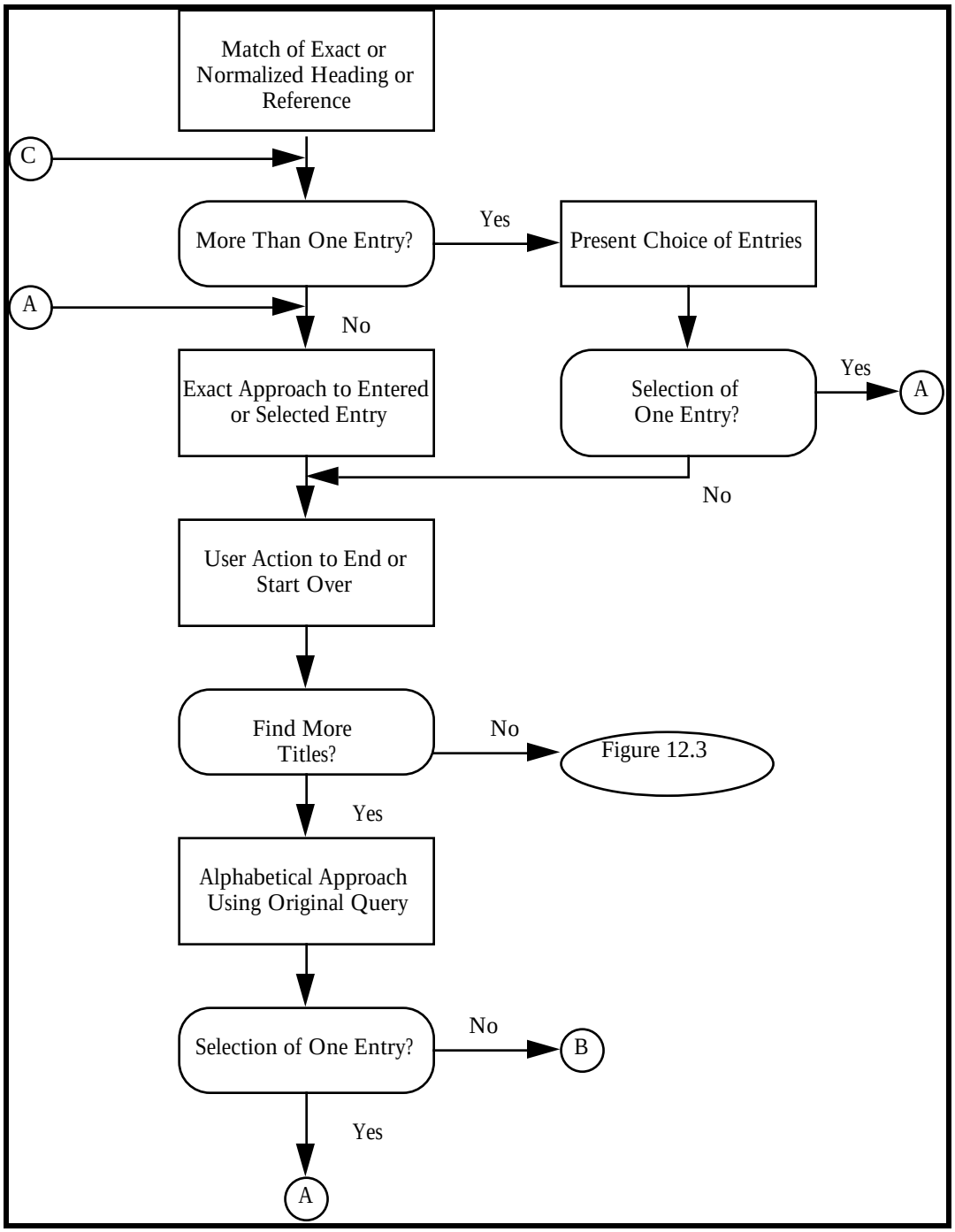

Figure 12.4A. Search tree for the exact approach

In figure 12.4A, systems invoke the exact approach. This approach gives users options for browsing related terms, notes, subdivided forms of matched terms, and expanding the search. Users who want to expand their search are given the results of the

alphabetical approach. Figure 12.4B shows the search approaches featured in search expansions. One-word queries are submitted to the keyword-in-main-heading search. Queries exceeding one word are submitted to the remaining controlled vocabulary searches, i.e., keyword-in-subdivided- and keyword-in-main-heading searches, and free-text searches, i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches.

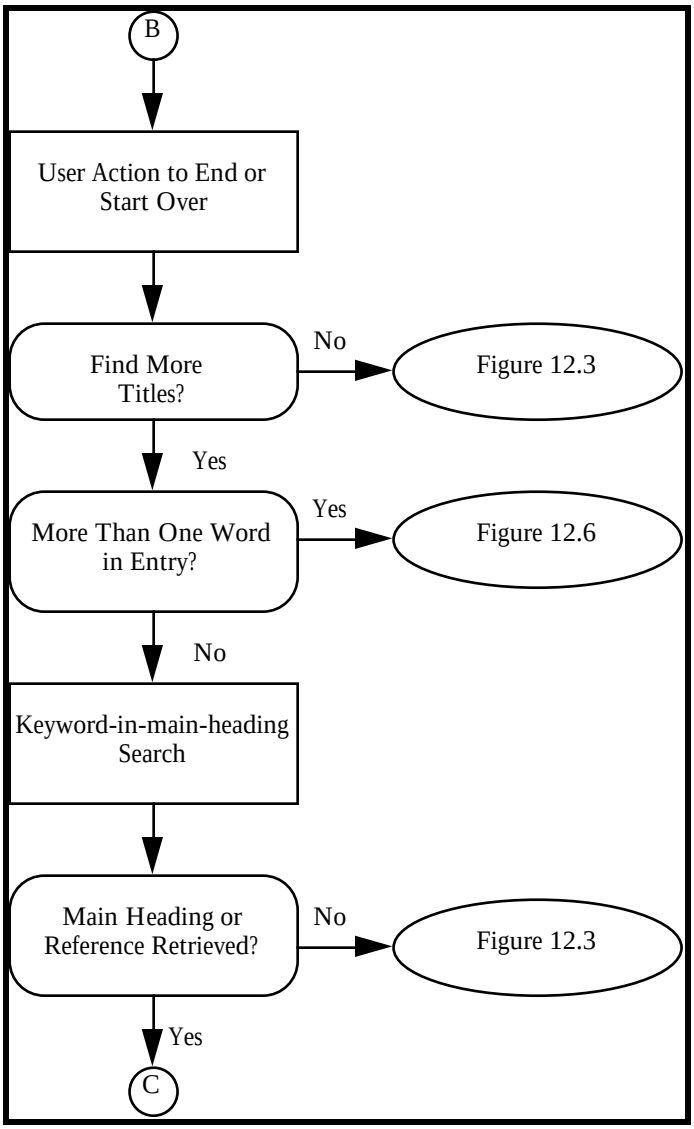

Figure 12.4B. Search tree for the exact approach (contd.)

12.4.2 Search Tree for One-word Queries

The redesigned search tree for one-word queries contains a few changes from the original search tree for one-word queries: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names (figure 12.3), (3) invoking the keyword-in-record search, (4) invoking relevance feedback following title-keyword and keyword-in-record searches, and (5) invoking the alphabetical approach as the search type "of last resort." The search tree for one-word queries is given in figures 12.5A and 12.5B.

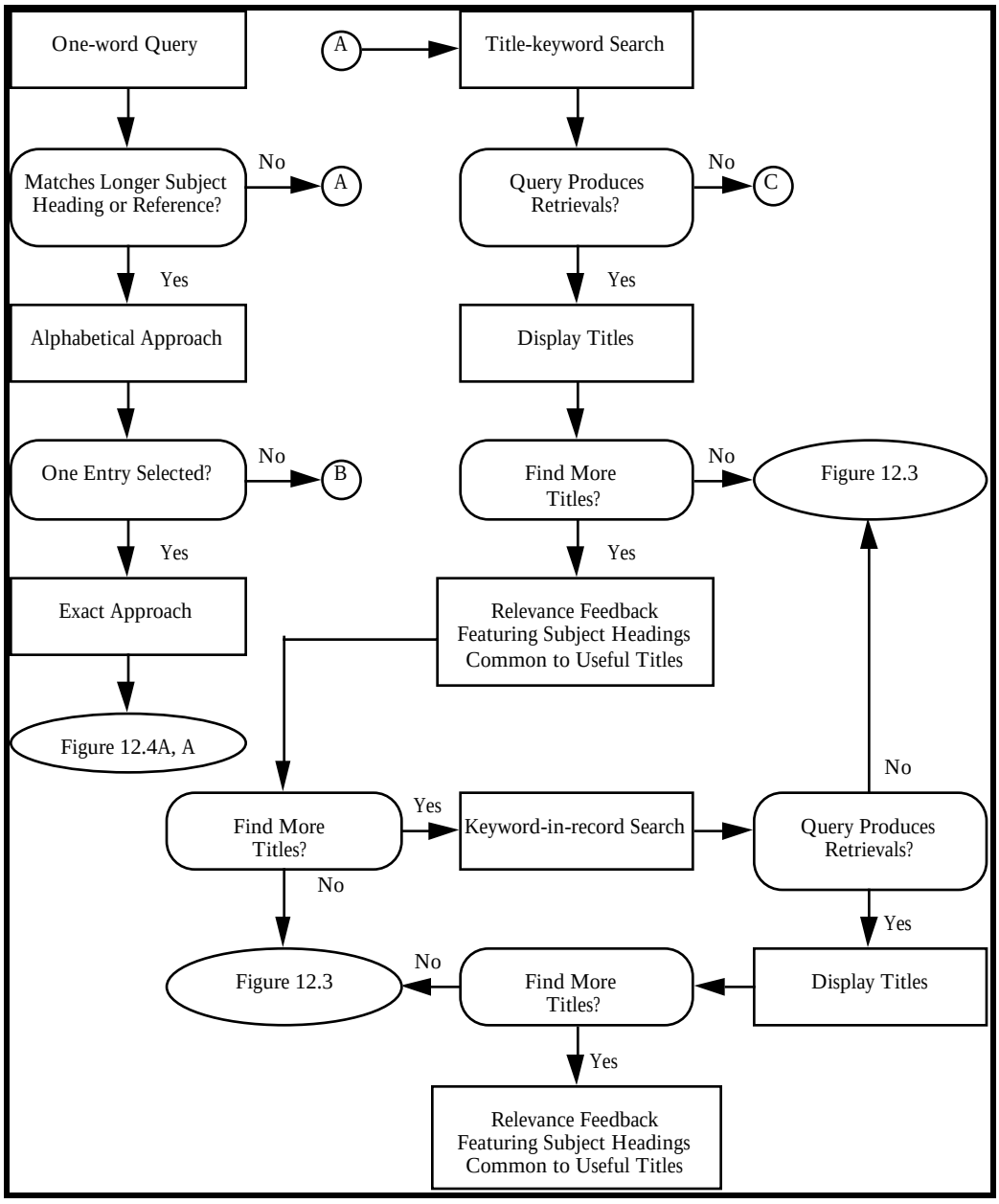

Figure 12.5A. Redesigned search tree for one-word queries

Figure 12.5A shows system actions that test whether one-word queries meet the criteria for invoking the alphabetical approach. Queries that fail are submitted to title-keyword searches. If they are successful, systems retrieve titles, and, at the conclusion of the title display, they ask users whether they want to retrieve additional titles. Systems respond to users who give positive responses with the results of searches for subject headings common to several retrieved titles that users rated useful. Systems can continue to find additional titles through keyword-in-record searches and relevance feedback

based on keyword-in-record search results. Of course, relevance feedback assumes that systems collect relevance assessments during the display of retrieved titles.

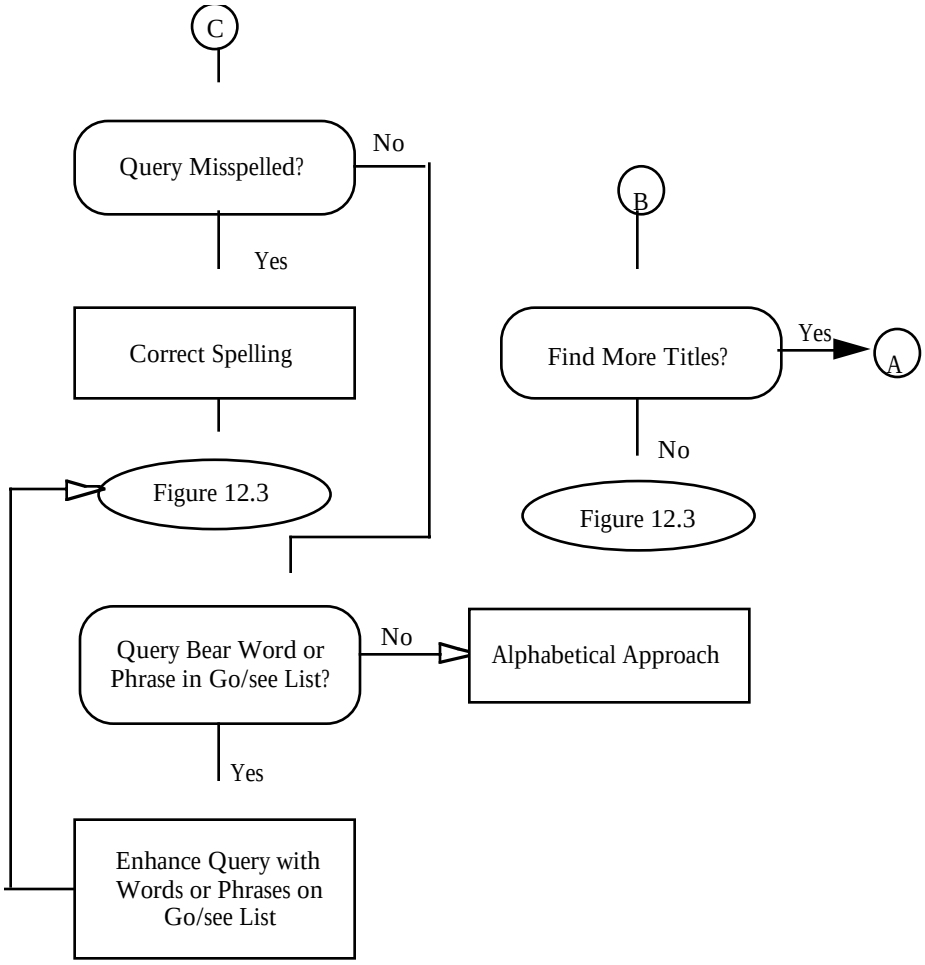

Figure 12.5B. Redesigned search tree for one-word queries (contd.)

The original search trees for one-word queries did not include keyword-in-record searches (see figure 4.3). This search type is included in the final version of the search trees for one-word queries. In view of the low levels of perseverance experimental online catalog users exhibited, few users will reach the results of keyword-in-record searches for one-word queries unless they continually expand search results. On occasion, however, keyword-in-record searches might be fruitful in retrieving useful titles for users whose queries fail to retrieve titles through exact, alphabetical, and title-keyword searches. For example, a contents note or summary field of a bibliographic record might bear a word matching the query. The only way for systems to retrieve this record is through a keyword-in-record search.

Figure 12.5B shows the operations that systems perform on queries that fail to produce retrievals in title-keyword searches. They first check user queries for possible misspellings. Systems could assist users in correcting spelling errors using spelling-correction routines similar to such routines for word processing programs in which they suggest alternate spellings for misspelled words. If users fail to correct spelling errors, systems could check queries against go/see lists to determine if queries bear listed words or phrases. They would then enhance queries with words and phrases from go/see lists and start at the initial search tree to find matches. Section 12.1.5 discussed candidate types of terms and phrases for inclusions in go/see lists. Examples are one-word/two-word variations of concepts and noun/adjective pairs. If systems enhance queries with words and phrases from go/see lists, search trees start begin with the initial search tree and try to effect matches of the system-enhanced user query.

The alphabetical approach has been added to the search tree for one-word queries as the search type "of last resort." If all other subject searches fail to produce retrievals, systems should respond with the results of the alphabetical approach, i.e., a display of subject headings and *see* references in the alphabetical neighborhood of the user-entered term. If, however, the user query matched the criteria for an alphabetical search following the exact approach, the search type "of last resort" is the keyword-in-record search.

12.4.3 One-word Search Tree for Small Databases

Users who search small databases that feature search trees might find expanding their search somewhat tiresome. Expanding could be especially tiresome in the event that expanding produces few new titles that users have not retrieved in the results of previous subject searches. System designers could consolidate title-keyword and keyword-in-record searches into a single keyword search. The search tree for one-word queries would feature three subject searching approaches for small databases: (1) exact, (2) alphabetical, and (3) keyword-in-record searches. The latter should feature relevance feedback to find additional titles. Relevance feedback would involve an exact search for subject headings common to titles judged useful by system users.

12.4.4 Search Trees for Multi-word Queries

The redesigned search tree for multi-word queries features the following enhancements: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names (figure 12.3), (3) performing relevance feedback based on useful retrievals in free-text searches (i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches), (4) invoking stemming, (5) invoking the best-match approach, and (6) invoking the alphabetical approach as the search type "of last resort." The

redesigned tree is shown in figures 12.6A–C. The search tree for multi-word queries begins with system efforts to invoke the alphabetical approach (figure 12.6A).

If user queries fail to meet the criteria for the alphabetical approach, systems perform keyword-in-record searches for the individual words of user queries. If systems fail to produce retrievals for one or more query words, they present the words to users for spelling correction (figure 12.6C). Systems could assist users in correcting spelling errors using spelling-correction routines similar to such routines for word processing programs in which they suggest alternate spellings for misspelled words. If users fail to correct spelling errors, systems could check queries against go/see lists to determine if queries bear listed words or phrases (figure 12.6C). They then enhance queries with words and phrases from go/see lists and start at the initial search tree to find matches.

If the individual words in user queries produce retrievals in keyword-in-record searches, systems do not immediately show users the results (figure 12.6A). Instead, they pass queries onto other controlled vocabulary searches — the keyword-in-main-heading and keyword-in-subdivided-heading searches (figure 12.6A). If these searches produce too few or no retrievals, systems submit queries to free-text approaches (figure 12.6B).

Free-text approaches begin with the title-keyword search (figure 12.6B). If searchers want to find additional titles, systems can invoke relevance feedback in which they retrieve additional titles based on searches for subject headings common to titles users rated useful in title-keyword searches. If title-keyword searches fail to produce retrievals, systems can try keyword searches of subject heading fields and keyword-in-record searches. Following successful matches, systems can invoke relevance feedback using subject headings common to titles users rate useful to find additional titles. Of course, relevance feedback assumes that systems collect relevance assessments during the display of retrieved titles.

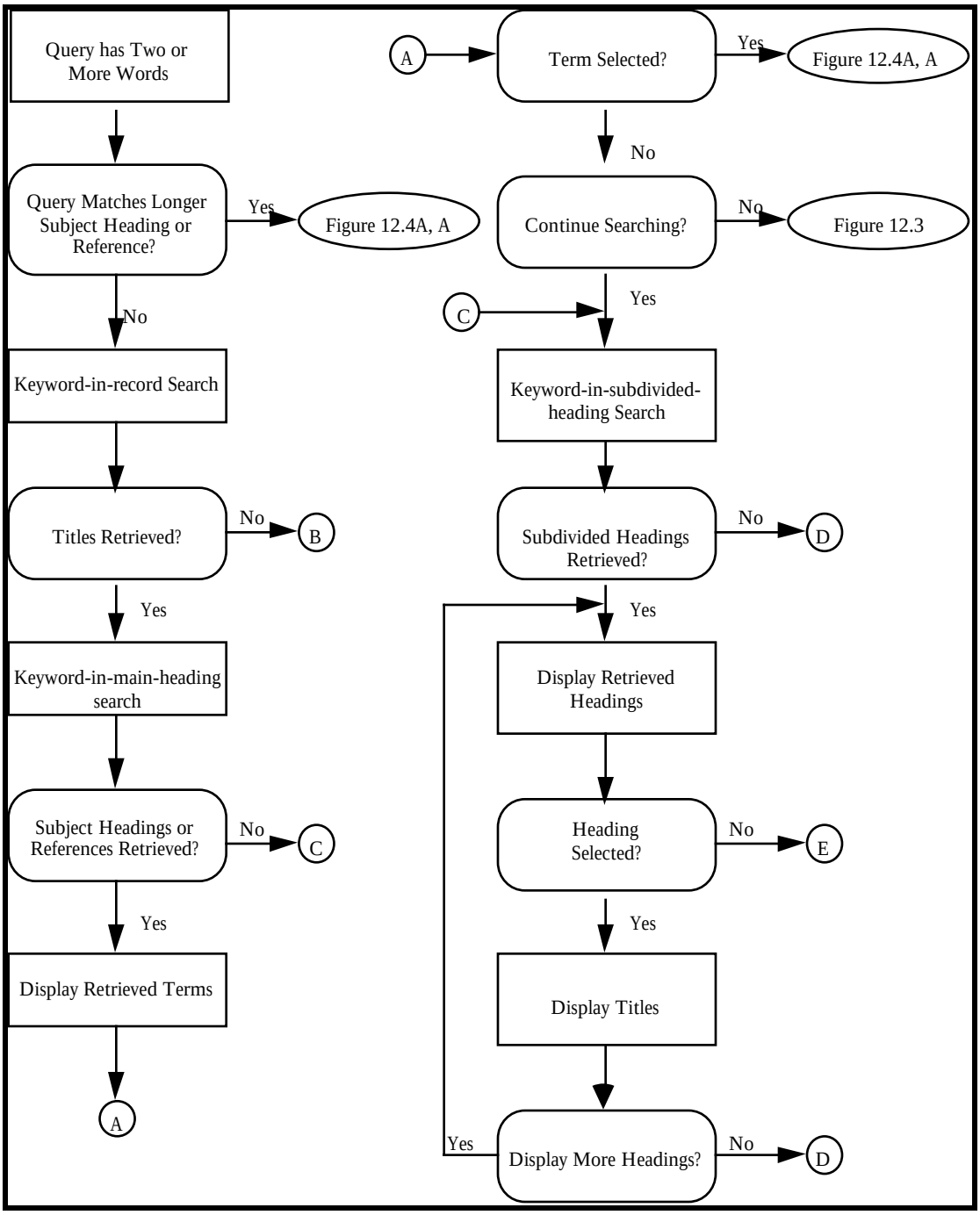

Figure 12.6A. Redesigned search tree for multi-word queries

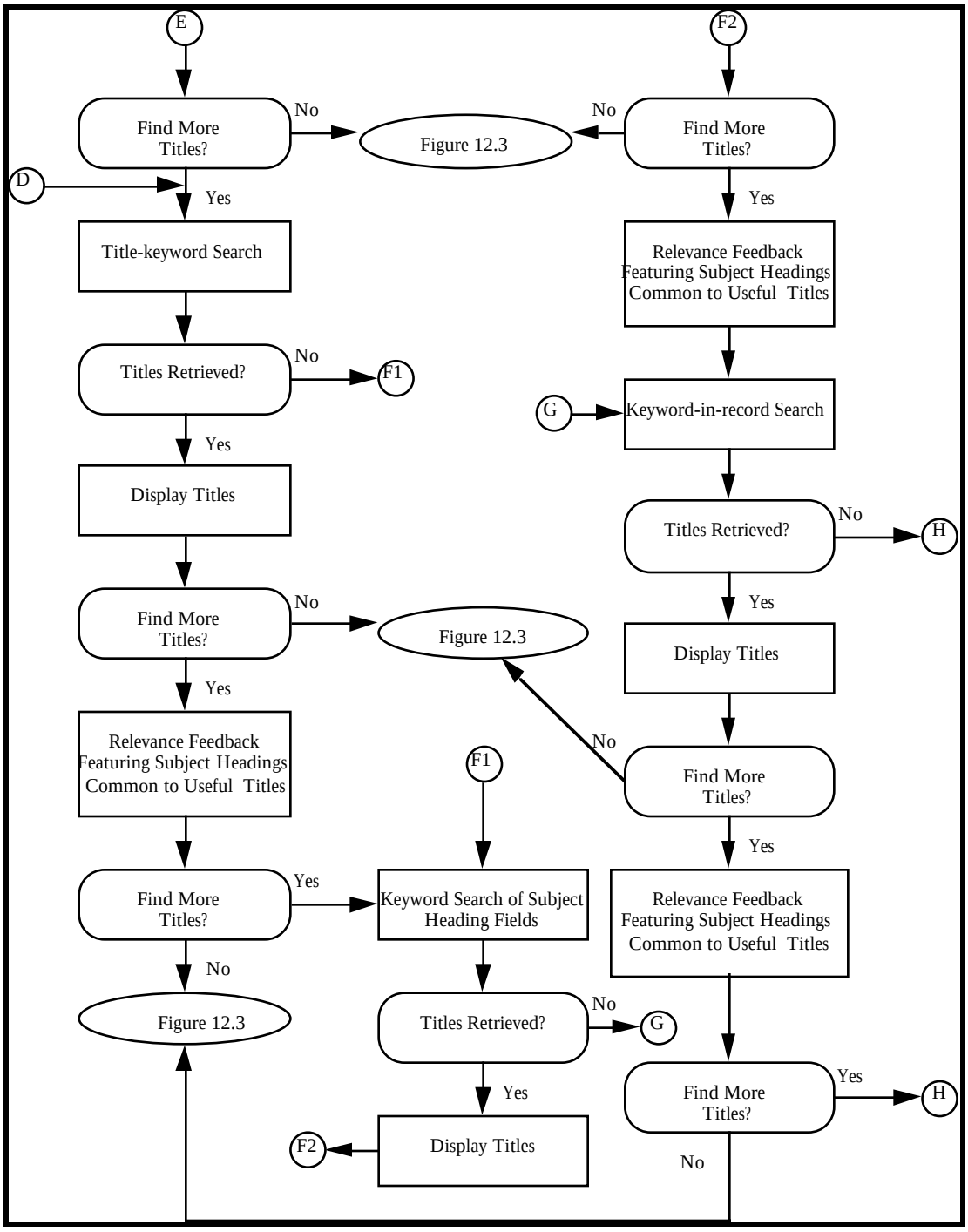

Figure 12.6B. Redesigned search tree for multi-word queries (contd.)

Following the keyword-in-record search are two new approaches (figure 12.6C). One approach involves stemming of query words. The order of indexes that systems would search for stemmed query words are exact, keyword-in-main-heading, keyword-in-subdivided-heading, title-keyword, keyword in subject heading fields, and keyword-in-record. If systems are unsuccessful making matches of user queries, they should continue searching with the best-match approach.

The best-match approach would feature stemming, weighted-term probabilistic retrieval, and output ranking. Systems would perform searches on a best match, combinatorial basis. A match on fewer than all query-word stems may retrieve titles but titles that have all the query-word stems would be displayed first.

Query-word stems would be weighted so that frequently-occurring stems would get a low "importance" weight compared to the weight assigned to rare word stems. A new weighting system should be devised for library cataloging databases because subject content is represented in so few fields of bibliographic records, i.e., title and subject heading fields. Systems should increase the weight given to frequently-occurring words in subject headings and subdivisions because matches on such words are likely to retrieve subject headings describing users' topics of interest or a facet of their topics of interest.

The alphabetical approach has been added to the search tree for multi-word queries as the search type "of last resort." If all other subject searches fail to produce retrievals, systems should respond with the results of the alphabetical approach, i.e., a display of subject headings and *see* references in the alphabetical neighborhood of the user-entered term.

12.4.5 Multi-word Search Tree for Small Databases

Users who search small databases that feature search trees might find expanding their search somewhat tiresome. Expanding might be especially tiresome in the event that expanding produces few new titles that users have not retrieved in the results of previous subject searches.

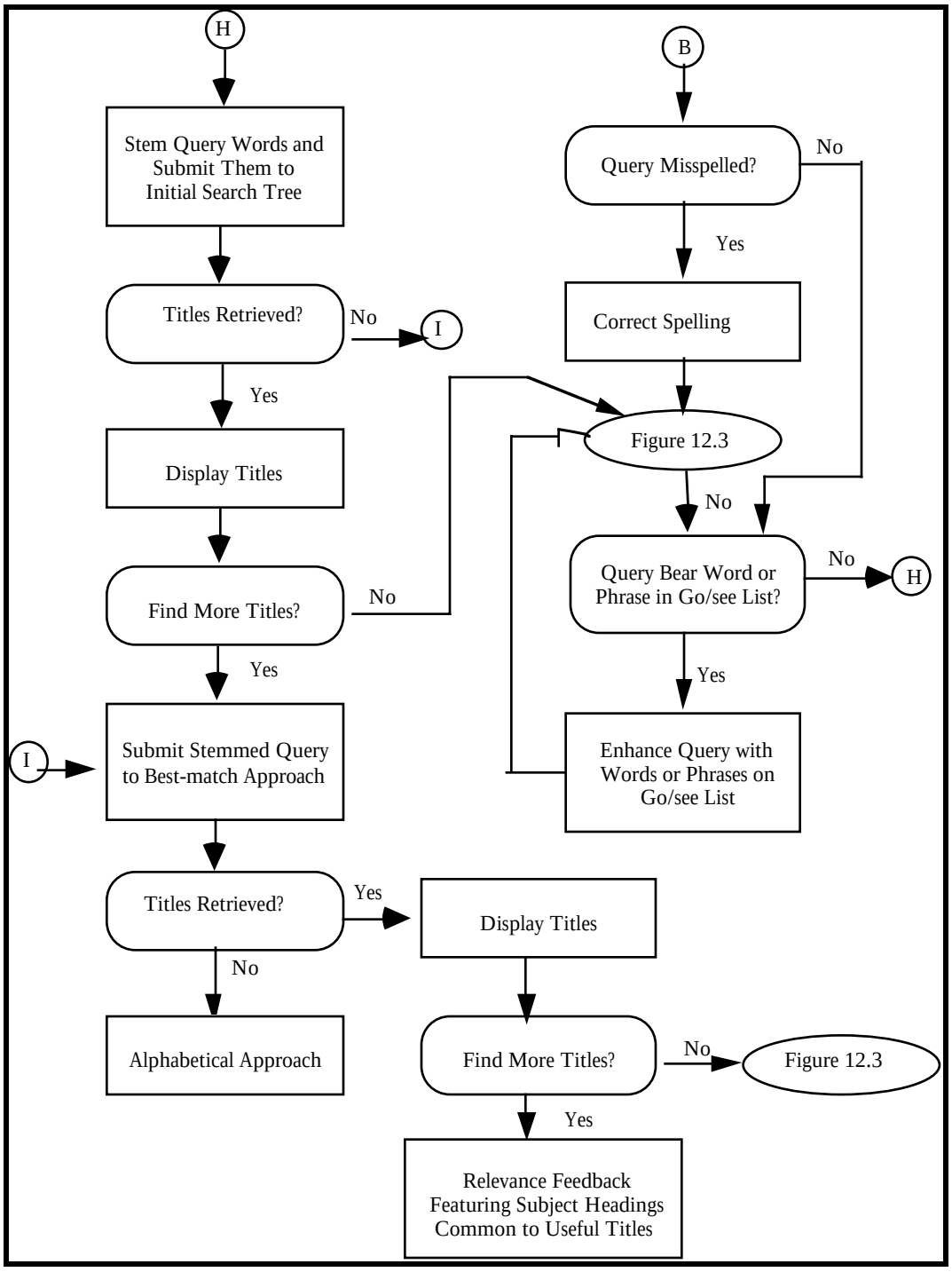

Figure 12.6C. Redesigned search tree for multi-word queries (contd.)

System designers could consolidate the results of the following searches: (1) keyword-in-main and keyword-in-subdivided-heading searches, (2) title-keyword, keyword in

subject heading fields, and keyword-in-record searches, and (3) stemming and best-match approach. The search tree for multi-word queries would feature five subject searching approaches for small databases: (1) exact, (2) alphabetical, (3) keyword-in-heading, (4) keyword-in-record, and (5) best match. Keyword-in-record and best-match searches would still feature relevance feedback to enable users to find additional titles on their topics of interest. Relevance feedback would involve an exact and keyword searches for subject headings common to titles judged useful by system users in keyword-in-record and best-match searches.

12.4.6 Search Trees for Subject Queries for Personal Names

The search tree for personal-name subject queries remains unchanged from the original conception (see Table 4.1). The search trees should be accompanied by pop-up windows that prompt users for the elements of their personal-name queries. This prompting enables systems to differentiate between name and topic elements of queries and choose to ignore certain elements in their attempts to retrieve titles. More research is needed to ensure that users enter the correct elements in various pop-up windows but this should not deter system designers from implementing the search trees and prompting mechanisms in operational systems.

The first step is for systems to single out queries containing topics. These queries are submitted to the keyword-in-subdivided-heading search followed by the keyword-in-record search in the hopes of finding headings and bibliographic records containing both name and topic elements. If systems fail to find both elements, they omit the topic element from the query and continue searching through the alphabetical approach. Personal name queries consisting exclusively of name elements are submitted to the alphabetical approach only.

12.5 Chapter Summary

The analysis of online retrieval test data demonstrated that the search trees were more effective in selecting a subject searching approach that would produce useful information for the subjects users seek than users would select on their own. However, the search trees require enhancements to enable them to respond with useful retrievals to especially difficult user queries, that is, those for which the existing set of search trees produce no retrievals.

Chapter 12 focuses on the redesign of search trees for subject searching in online catalogs. Two problems — spelling and vocabulary problems — were greatly

responsible for forcing a redesign of the search trees (sections 12.2 and 12.3). The analysis of spelling errors in user queries demonstrated that online catalogs should feature computer-assisted spelling correction routines — just like word processing programs — to assist end users in identifying and correcting their spelling errors. Solving vocabulary problems would require several enhancements to search tree configurations: (1) eliminating punctuation in queries including double posting hyphenated words in subject headings and title fields, (2) creating a go/see list to treat words and phrases as equivalent terms during retrieval, (3) automatic stemming of queries to effect retrievals in exact and the various keyword searches, and (4) best-match approach to retrieve titles for the most difficult user queries that traditional searches — exact, alphabetical, and keyword — cannot handle.

The only change to the initial search tree from the original design was returning users whose searches failed to produce retrievals to the question about personal names (see figure 12.3).

The search tree for the exact approach remained unchanged. Systems respond to user queries for subjects generally that match exact or normalized forms of controlled vocabulary terms with the exact approach (figures 12.4A and 12.4B).

The redesigned search tree for one-word queries contained a few changes from the original search tree for one-word queries: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names, (3) invoking the keyword-in-record search, (4) invoking relevance feedback following title-keyword and keyword-in-record searches, and (5) invoking the alphabetical approach as the search type “of last resort.” The search tree for one-word queries is given in figures 12.5A and 12.5B.

The redesigned search tree for multi-word queries featured the following enhancements: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names, (3) performing relevance feedback based on useful retrievals in free-text searches (i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches), (4) invoking stemming, (5) invoking the best-match approach, and (6) invoking the alphabetical approach as the search type “of last resort.” The redesigned tree is shown in figures 12.6A–C.

The search tree for personal-name subject queries remained unchanged from the original conception (Table 4.1). The first step is for systems to single out queries containing topics. These queries are submitted to the keyword-in-subdivided-heading search followed by the keyword-in-record search in the hopes of finding headings and bibliographic records containing both name and topic elements. If systems fail to find both elements, they omit the topic element from the query and continue searching

through the alphabetical approach. Personal name queries consisting exclusively of name elements are submitted to the alphabetical approach only.

Chapter 12 also contains recommendations for abridged search trees for subject searching in small databases (sections 12.4.3 and 12.4.5). Users who search small databases that feature search trees might find expanding their search somewhat tiresome. Abridged search trees for one-word queries feature only three subject searching approaches: (1) exact, (2) alphabetical, and (3) keyword-in-record searches with relevance feedback. The search tree for multi-word queries would feature five subject searching approaches for small databases: (1) exact, (2) alphabetical, (3) keyword-in-heading, (4) keyword-in-record with relevance feedback, and (5) best match with relevance feedback.

References

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. Using subject headings for online *retrieval: Theory, practice, and potential*. San Diego: Academic Press.

Walker, Stephen, and Richard M. Jones. 1987. *Improving subject retrieval in online catalogues; 1. Stemming, automatic spelling correction and cross-reference tables*. London: British Library. British Library Research Paper no. 24.

13 Highlights of Project Activities and Findings

13.1 Project Overview

This report describes the activities, analyses, findings, and recommendations of a research project that tests a new design for subject access to online catalogs. The foundation for the new design were findings from an empirical study of the subject terms users entered into online catalogs (Drabenstott and Vizine-Goetz 1994). The new design required that online catalogs have a wide range of subject searching capabilities and search trees to govern the system's selection of searching capabilities in response to user queries. Search trees hold much promise for assuming the burden of determining which subject searching approach is likely to produce useful information in response to user queries.

Project objectives were to:

1. Demonstrate subject searching in an experimental online catalog with a wide range of subject searching capabilities and with search trees to govern system selection of a particular capability in response to user queries.
2. Test the retrieval effectiveness of the experimental online catalog with search trees by comparing its performance with the performance of an experimental online catalog in which subject searching approaches are assigned at random.
3. Evaluate the demonstration and test results of retrieval effectiveness and disseminate the research findings through publications in the professional literature.

The research project was composed of twelve activities grouped into three phases (figure 1.1). Seven activities (1–6, 9) were required in the system-development phase. This phase featured the development of the ASTUTE experimental online catalog with a wide range of subject searching functionality and search trees to govern the system's

selection of a subject searching approach in response to user queries. The databases of the experimental online catalog contained bibliographic records in computer science and technology from Mardigian Library of the University of Michigan-Dearborn (UM-D) or bibliographic records in American history from Lilly Library of Earlham College, and subject authority records from the Library of Congress' subject authority file. Three activities (7–8, 10) were required in the system testing phase. Data collection procedures and instruments were developed, pretested, and revised. Instruments were administered in online retrieval experiments with library patrons and staff at Mardigian and Lilly Libraries of UM-D and Earlham College to answer the following three research questions:

1. Do search trees improve the search performance of subject searchers at the online catalog?
2. Do subject searchers prefer an online catalog that controls system responses and searching approaches to an online catalog without such controls?
3. What are the characteristics of queries that cannot be answered using the ASTUTE experimental online catalog?

Two activities (11–12) made up the evaluation phase of the project. Collected data were described, analyzed, interpreted, and synthesized in a final report to the Department of Education. The ASTUTE Project Team also plans to disseminate findings in journals and conference proceedings to the library and information science communities.

13.2 Development of the ASTUTE Experimental Online Catalog

13.2.1 Computer Equipment

ASTUTE (A Search Tree Underlying the Experiment) was programmed on a stand-alone Gateway 2000 486, 33 MHz, IBM-compatible microcomputer, with 8 megabytes of RAM and a VGA color monitor (section 3.2). The operating system was MS-DOS version 5.0. A dot-matrix printer and a mouse were attached to the microcomputer for use by ASTUTE project staff during development work and during online retrieval tests with end users and librarians.

When ASTUTE was installed in the libraries of UM-D and Earlham College, the system did not require a network connection. It resided entirely on the Gateway microcomputer. Participating library staff monitored the system, performed daily system backups, and periodically used the microcomputer equipment and network

connections in their institutions to transmit transaction log files of end-user search activity to the ASTUTE project team in Ann Arbor.

13.2.2 Database Development

The databases of the ASTUTE experimental online catalog were created from two data sources: (1) machine-readable cataloging (MARC) records for bibliographic data from the two participating libraries (section 3.3), and (2) USMARC records for subject authority data from the compact disc-based product CD/MARC Subjects distributed by the Library of Congress (section 3.4). Bibliographic records from the University of Michigan-Dearborn and Earlham College were limited to computer science and technology, and American history, respectively. Subject authority records were limited to subject headings assigned to bibliographic records.

The team made a few changes to bibliographic data: (1) deleting diacritics in Earlham bibliographic records (section 3.3.3.1), and (2) correcting erroneous subject headings (section 3.3.3.2). The few enhancements the team made to bibliographic data were intended to enhance subject access: (1) identifying and coding form subdivisions to increase the number of broad categories in exact search from three to four (section 3.3.4.1), and (2) converting control-field codes to English-language equivalent terms to increase the searchable dictionary in keyword-in-record searches (section 3.3.4.2). The team deleted diacritics in subject authority records (section 3.4.2) and enhanced records with narrower terms (section 3.4.3).

13.2.3 Search Trees and System Functionality

The ASTUTE experimental online catalog was actually composed of two online catalogs: (1) the Blue System, in which search trees governed the system's selection of a subject searching capability, and (2) the Pinstripe System, in which the system selected a subject searching capability randomly. These systems were purposely designed to be very much alike to focus the attention of library patrons and staff on the retrieval of useful information in response to their queries. The Blue and Pinstripe Systems had virtually the same interfaces, and they accessed the same bibliographic and authority databases. Except for the Blue System's enhancement with the search trees, the two systems and their capabilities were very much the same.

The Blue and Pinstripe Test Systems used different criteria to choose subject searching approaches in response to user queries. Search trees governed the Blue System's selection of subject searching approaches (sections 4.2–4.5). Search trees were defined as a set of paths with branches or choices that enabled systems to carry out the most sensible search approach at each stage of the search (Mitev, Venner, and Walker 1985). The search

trees incorporated into the Blue System emanated from the findings of an empirical study of user queries (Drabenstott and Vizine-Goetz 1994). The Pinstripe System's selection of a subject searching approaches was governed by a random selection algorithm (section 5.10).

Search trees preferred two-step subject searching approaches that enlisted the catalog's controlled vocabulary, i.e., exact, alphabetical, and keyword-in-heading search approaches, before free-text approaches, i.e., title-keyword, keyword searches of subject heading fields, and keyword-in-record searches. If the former approaches supplied users with appropriate controlled vocabulary terms for expressing their topics of interest, they provided users with several opportunities to refine their searches using related terms, subdivided forms of the matched or selecting heading, or controlled vocabulary terms in close alphabetical proximity to the subject queries users entered into online catalogs.

The Blue System featured four controlled vocabulary approaches: (1) exact (subjects only), (2) alphabetical (subjects and personal names), (3) keyword-in-subdivided-heading (subjects only), and (4) keyword-in-main-heading (subjects only) searches. It featured three free-text approaches: (1) title-keyword (subjects only), (2) keyword in subject heading fields (subjects only), and (3) keyword-in-record (subjects and personal names) searches.

The Pinstripe System's selection of a subject searching approaches was governed by a random selection algorithm (section 5.10). This algorithm chose between three different subject searching searches: (1) alphabetical (subjects only), (2) keyword-in-subdivided-heading (subjects only), and (3) keyword-in-record (subjects and personal names).

Both Blue and Pinstripe Systems prompted users for the last name, first name, and topic elements of their subject queries for personal names (section 5.9.2). The Pinstripe System always responded to such queries with the keyword-in-record search. The Blue System preferred the keyword-in-record search for personal-name queries. If it was unsuccessful retrieving titles, it discarded the topic element and submitted personal-name queries to the alphabetical approach.

13.3 Methods Used to Test the New Subject Access Design

The methods used in this project to compare the performance of the Blue and Pinstripe Systems were based on the Comparison Search Experiment devised by Siegel et al. (1983) in their evaluation of two prototype online catalogs at the National Library of Medicine (NLM) (section 6.2). Comparative approaches to system evaluation like

Siegel's and several modeled after Siegel's experiment have used human intermediaries to collect information about searchers and their searching experiences. Such approaches are expensive and time-consuming for several reasons (section 6.3). For example, interviewers must be present to monitor online searches, ask users questions, and record their answers. The project director used a comparative approach to system evaluation in this project that did not use interviewers or human intermediaries at all. The experimental online catalog performed and/or recorded all activity connected with data collection — searcher recruitment, questionnaire administration, search logging, and searcher relevance assessments.

In the Comparison Search Experiment, the Blue and Pinstripe Systems encouraged library patrons and staff to perform online subject searches in both systems (section 6.4.1). In introductory screens, the systems informed users that they were participating in an experiment and allowed them to graciously exit the system if they did not want to take part. Prior to conducting subject searches, the systems asked participants to respond to three pre-search questions that collected demographic information on participants (section 6.4.2). During their searches, the systems recorded user activity and user relevance assessments for displayed bibliographic records to a time-stamped transaction log (section 6.6 and Appendix F). The systems asked participants to answer eleven questions in a post-search questionnaire that compared the performance of the Blue and Pinstripe Test Systems (section 6.4.4 and Appendix C). Thus, the retrieval tests yielded quantitative comparative data about the systems' capabilities and effectiveness, and about users' system preference. After completing the post-search questionnaire, library staff were asked additional questions that prompted them to make comparisons between the two systems with respect to the systems' ease of use, subject searching capabilities, and performance (section 6.5 and Appendix E). These qualitative data were combined with quantitative data from patron searches to increase our understanding of system performance.

13.4 Comparison Search Experiment Participants

Before the researchers could submit Comparison Search administrations to the data analysis, they had to manually review individual search administrations to determine whether patron search topics were in scope and whether patron responses to pre- and post-search questionnaires were valid.

The three installations of the experimental online catalog lasted for five- to fifteen-week periods at the two participating libraries. Data collected in the first of two data-collection periods at UM-D were discarded because of a system error that corrupted data for over two-fifths of the search administrations (section 7.4). The system error

was introduced by the unexpected user action of repeatedly turning the microcomputer on and off. The programming team made changes to the experimental online catalog to ensure that the data would not become corrupted in later system installations.

In the second, five-week data-collection period at UM-D, the experimental online catalog administered a total of 826 search experiments (section 7.5). In the fifteen-week data-collection period at Earlham, the experimental online catalog administered a total of 238 search experiments (section 7.6).

The manual review of searches revealed that about 50% of search administrations at UM-D and Earlham were usable (sections 7.5.2 and 7.6.2); however, the manual review of post-search responses revealed that large percentages (58% at UM-D, 51% at Earlham) of search experiments were less-than-full administrations. That is, users completed a search in one system, both systems, and/or failed to answer the post-search questionnaire. Since monitors were not present to encourage patrons to complete the entire Comparison Search Experiment, the researchers expected patrons to leave before they finished the full administration of the Comparison Search Experiment. The manual review process was essential because it identified full and less-than-full search administrations which were, on occasion, handled differently in the statistical analysis of search results. Only searches with usable queries were submitted to statistical analyses and failure analyses (chapters 8–12).

Unusable administrations were divided into categories for queries that were out of scope, for play, meaningless input, known-item searches, and other unusable searches (e.g., sex terms, expletives) (sections 7.5.3 and 7.6.3). Much of sections 7.5 and 7.6 was given to discussions of unusable search administrations — systems searched, titles retrieved, relevance assessments, etc. Discouraged with the preponderance of positive relevance assessments users gave to displayed titles in unusable search administrations, the researchers made suggestions for redesigning the procedures for collecting relevance assessments in future experimental and operational systems (section 7.7). They also made suggestions for redesigning system-administered post-search questionnaires to minimize the collection of unreliable data in future experimental systems (section 7.8).

13.5 Characteristics of System Users and Searches

13.5.1 System Users

The majority of users at both UM-D and Earlham were searching ASTUTE for the first time (section 8.2). The majority were also daily users of computer systems. Large percentages of UM-D respondents were computer science or engineering majors

(29.1% and 46.1%, respectively). Thus, the majority of UM-D respondents were performing searches for topics in their major field of study. This was not true of respondents at Earlham College where ASTUTE's database focused on American history. Only one-quarter of Earlham respondents were history majors; consequently, the majority of Earlham respondents were not performing searches for topics in their major field of study.

Searchers did not spend much time using the ASTUTE experimental online catalog (section 8.3). They spent less than a minute completing the pre-search questionnaire. Time spent searching was quite low for search administrations bearing suspect post-search questionnaires — a little more than three minutes at UM-D and two minutes at Earlham. Search administrations were the lengthiest for complete search administrations — over four minutes at both UM-D and Earlham. Completing post-search questionnaires took less than a minute in search administrations with suspect questionnaire responses and took a minute and a half in full search administrations.

13.5.2 Searches

ASTUTE was designed to handle one query per search experiment administration (section 8.4). Thus, the average number of queries per search (1.2) was much lower than the average number in operational online catalogs. A statistical test demonstrated that the length of queries had a significant effect on the ability of the experimental online catalog to make matches of its subject vocabulary (Table 8.5). Long user queries (three or more words) were less likely to make matches than short queries (two or fewer words).

13.5.3 Matches of User Queries

The first comparison of system performance was given in an analysis of system responses that designated matches of user queries (section 8.5.1). Matches gave users opportunities to pursue the results of subject searching approaches. Generally, the Blue System was likely to provide more opportunities for guiding users to useful information on their topics of interest than the Pinstripe System provided (Table 8.6).

Providing users the opportunity to display titles did not guarantee that users would, especially in the case of controlled vocabulary approaches, choose a seemingly relevant subject heading from a list that would take them one step closer to retrieving titles connected with the subject heading. Furthermore, it did not guarantee that users would browse more than the first title displayed in free-text approaches. Users seldom displayed all the titles they retrieved, in fact, the statistics in Table 8.8 showed that they displayed about a fifth of the titles they retrieved. At UM-D, the Pinstripe System

retrieved $2\,^1/_2$ times as many titles as the Blue System. High-posted searches in the Pinstripe System were likely to come from the keyword-in-record search; this search made retrievals based on matches in all subject-rich fields. Pinstripe-system users at UM-D displayed only about an eighth of the titles they retrieved.

13.5.4 Resolving Conflicting Relevance Assessments

At UM-D, Blue- and Pinstripe-system users displayed a total of 470 titles more than once. Of this total, 82.8% of titles were displayed twice in across systems, 11.5% were displayed three times across systems, and 5.7% were displayed more than twice across systems. At Earlham, Blue- and Pinstripe-system users displayed a total of 176 titles more than once. Of this total at Earlham, 80.1% of titles were displayed twice across systems, 12.5% were displayed three times across systems, and 7.4% were displayed more than three times across systems. The researchers reviewed relevance assessments for the same titles and noticed that the more users displayed the same titles, the more likely they were to give positive relevance assessments to titles, i.e., "very useful" or "possibly useful" ratings. After much discussion among team members and consultants, we decided to use the last relevant assessment that searchers gave to titles displayed more than once and to disregard other relevance assessments for the same title. The resolution of such conflicts did not have a significant effect on sets of rated titles (Table 8.10). Resolution was necessary, however, to submit the data to statistical procedures to measure precision and estimated recall.

13.6 Precision and Estimated Recall in the Blue and Pinstripe Systems

13.6.1 Precision Scores

At UM-D, precision scores averaged 64% in the Blue System and 58% in the Pinstripe System (section 9.2.1). In full search administrations at UM-D (i.e., administrations in which users conducted searches in both Blue and Pinstripe Systems and completed a post-search questionnaire), precision scores in the Blue System were significantly higher than precision scores in the Pinstripe System (figure 9.1. and Table 9.1). The result was exactly the opposite at Earlham. That is, precision scores in the Pinstripe System were significantly higher than precision scores in the Blue System (figure 9.2. and Table 9.2).

We extended the analysis of precision scores to comparisons for controlled vocabulary and free-text searches because the Blue System's search trees favored controlled vocabulary over free-text searches in response to user queries (section 9.2.2). We

expected the precision of titles retrieved through controlled vocabulary searches to exceed the precision of titles retrieved through free-text searches. Such results would support the design of search trees in the experimental online catalog.

In UM-D's Blue and Pinstripe Systems, precision scores for controlled vocabulary searches exceeded such scores for free-text searches (figures 9.3 and 9.4). Furthermore, the differences between scores were significant (Tables 9.3 and 9.4). Especially large differences between precision scores were registered for full search administrations. In Earlham's Blue System, precision scores for controlled vocabulary approaches exceeded such scores for free-text approaches (figure 9.5). Again, the difference between scores was especially large for full administrations. There was an insufficient number of searches for statistical comparisons of precision in Earlham's Pinstripe System; however, overall precision scores for controlled vocabulary approaches were significantly higher than such scores for free-text approaches (Table 9.6). These findings about precision and search approaches support the design of search trees for searching bibliographic databases that favor controlled vocabulary over free-text approaches.

13.6.2 Recall Scores

Overall, recall scores in UM-D's experimental online catalog were about the same (section 9.3). Recall in the Blue System was 53.4% and recall in the Pinstripe System was 55.4% (figure 9.6). Recall scores in the two search administrations flip flopped. That is, recall scores in the Blue System exceeded such scores in the Pinstripe System in full search administrations and vice versa in partial administrations. Type of search administration results were the same at Earlham — recall scores in the two search administrations flip flopped (figure 9.7). That is, recall scores in the Blue System exceeded such scores in the Pinstripe System in full search administrations and vice versa in partial administrations. Overall, recall in Earlham's Pinstripe System (61.8%) was somewhat higher than recall in the Blue System (50.8%). No statistical tests were performed on recall scores because the design of the comparison search experiment violated the t-test's principle of independence.

13.7 Post-search Questionnaire Responses Regarding System Performance

End-user responses to the post-search questionnaire were available for full search administrations (section 9.4). Generally, responses reflected statistical results of this type of search administration. That is, users rated the performance of the Blue System higher than the Pinstripe System in terms of retrieving greater numbers of useful titles

(figure 9.8), and search satisfaction (figure 9.9), and they preferred the Blue System over the Pinstripe System (figure 9.10).

The post-search questionnaire probed other factors — ease of system use, system efficiency in retrieving useful titles, system helpfulness in giving users new subject searching ideas — in the comparison of Blue and Pinstripe Systems. The majority of end users rated the Blue System or both systems highly in terms of the ease of giving instructions to the systems (figure 9.11), ease of getting instructions from the system (figure 9.12), system clarity (figure 9.13), and system efficiency in guiding them to useful information (figure 9.14).

At Earlham, users were a little less enthusiastic about the Blue System's helpfulness in giving new ideas for subject searching (figure 9.15). Since the design of the Blue System was intended to give users new subject searching ideas through exact-approach browsing and expanding opportunities, we were somewhat concerned about this response. The results of a full-scale failure analysis of experimental catalog searches revealed the reasons why users answered the question in this way and provided many areas for improving this approach (chapters 10–12).

13.8 Failure Analysis of User Searches

13.8.1 Characterizing Successful and Unsuccessful Searches

This study featured a failure analysis of end-user searches to determine why the experimental ASTUTE online catalog was successful or unsuccessful in retrieving useful titles. The project director, an expert searcher of information retrieval systems, repeated end-user searches that were successful and unsuccessful in retrieving useful titles to identify the reasons why they succeeded or failed.

The first step in the failure analysis of searches performed at the University of Michigan-Dearborn and Earlham College was to characterize searches that users performed in the Blue and/or Pinstripe Systems as successful or unsuccessful:

- Successful search — users rated one or more titles "very useful" in a search performed in the Blue or in Pinstripe System.
- Unsuccessful search — users did not rate one or more titles "very useful" in a search performed in the Blue or in Pinstripe System.
- Unsuccessful pairs of searches — users did not rate titles "very useful" in searches performed in the Blue *and* Pinstripe Systems for the *same* topic.

At both UM-D and Earlham (sections 10.2 and 10.3), about 44% of searches conducted in the Blue System and 37% to 43% of searches conducted in the Pinstripe System were successful. About 19% to 25% of searches conducted in the Blue System or Pinstripe System were unsuccessful. About a third of searches conducted in the Blue *and* Pinstripe Systems for the *same* topic were unsuccessful.

13.8.2 Reasons for Failure

The failure analysis of these searches demonstrated that the following problems adversely affected the retrieval of potentially useful information from the experimental online catalogs: (1) user perseverance, (2) specificity that was not represented in user queries, (3) conflicting relevance assessments, (4) database failure, (5) large retrievals, (6) user query vocabulary, (7) search approach failures, (8) navigational problems, and (9) spelling problems. Sections 10.2 and 10.3 described the failure analysis of searches conducted by users at UM-D and Earlham, respectively.

13.8.3 Reclassifying Successful Searches

Explanations and examples of searches placed in categories for perseverance, query specificity, and conflicting relevance assessments provided evidence that these searches were really successful searches in terms of leading users to titles on their topics of interest. With regard to user perseverance, the systems gave users ample opportunities for browsing terminology or displaying titles on their topics of interest but users terminated their searches. In searches marred by query specificity, users gave negative relevance assessments to the titles that seemed relevant to their queries. Their actions led us to conclude that the systems were successful in terms of retrieving titles on the topics users expressed in their queries but they were not successful in terms of retrieving titles that fulfilled the specifications of queries users had in their minds but did not make explicit in their queries. In the failure analysis, we also considered searches bearing conflicting relevance assessments as successful searches because users had given the titles positive assessments in the course of their searches.

The reclassification of searches placed into categories for perseverance, query specificity, and conflicting relevance assessments from unsuccessful to successful searches resulted in success rate of about 72% for Blue-system searches at UM-D and Earlham and a success rate of between 49% and 59% for Pinstripe-system searches at UM-D and Earlham (Tables 10.17 and 10.31). Clearly, the Blue System performed better than the Pinstripe System in terms of retrieving and displaying titles on the topics users made explicit in their queries (sections 10.2.3 and 10.3.3).

13.8.4 Associating Specific Failures with Subject Searching Approaches

Success in the Blue System, however, was marred by a large percentage of searches in which users did not persevere in terms of browsing subject terminology and displaying retrieved titles (sections 10.2.1.1, 10.2.2.1, 10.3.1.1, and 10.3.2.1). Almost ten percentage points separated the results of Blue- and Pinstripe-system searches for searches marred by perseverance (Tables 10.17 and 10.31). Many Blue-system searches affected by perseverance featured controlled vocabulary search approaches. Navigational problems were also linked to user perseverance and controlled vocabulary searches (sections 10.2.1.8, 10.2.2.8, and 10.3.2.8). Of the several suggestions made to improve controlled vocabulary searches, the important ones were:

- Reducing the response time for searching selected terms and creating exact-search option menus.
- Reserving exact-search option menus for summaries of large retrievals and seeking other methods for summaries of one or a few retrievals.

Success in the Pinstripe System was negatively affected by the system's random selection algorithm for subject searching approaches. Between ten and twenty percentage points separated the results of Blue- and Pinstripe-system searches for search approach failures (Tables 10.17 and 10.31). The keyword-in-subdivided-heading search was usually to blame for search approach failures; however, the alphabetical approach had a large share of the failures. Subject searching approach failures were rare in the Blue System, accounting for only six failures in over 450 searches. The Blue System was guided in its selection of a subject searching approach by the search trees. Thus, the data and analyses in this project provided evidence that search trees improved the search performance of subject searchers at the online catalog in terms of guiding them to subject headings and titles that described the topics they made explicit in their queries.

13.9 Failure Analysis of Subject Searching Approaches

13.9.1 Controlled Vocabulary Approaches

Controlled vocabulary searches were especially plentiful in the Blue System and accounted for almost two-thirds of the searches conducted in this system (section 11.2). About half (52.5%) of controlled vocabulary searches were successful in retrieving at least one title rated "very useful" by system users. Many (54.5%) of these successful searches were searches characterized by no browsing. That is, users displayed retrieved

titles and terminated their searches without browsing available options for broader or narrower terms, subdivided terms, notes, or expanding the search. Only 14.4% of successful controlled vocabulary searches featured extensive browsing (Table 11.5).

Frequent reasons for the failure of controlled vocabulary searches in the Blue System were perseverance and navigation problems (sections 11.2.4 and 11.2.5). Some of the searches in which perseverance was a problem could have been navigation problems. Based on the analysis of searches that failed due to perseverance and navigation problems, the following suggestions were made for improving the exact search:

- Reducing the complexity of the exact-approach main menu.
- Consolidating several exact-approach main menu options into fewer options based on numbers of retrievals.
- Including end users in the design and pretesting of the exact-approach main menu.
- Pretesting the wording of exact-approach main menu options with end users.
- Designing and implementing an easy, understandable backtracking option.

The keyword-in-subdivided-heading search was a controlled vocabulary search but it was treated separately from other controlled vocabulary searches because it was independent of the exact approach (section 11.3). Two-thirds (29 of 44) of successful keyword-in-subdivided-heading searches featured no browsing (section 11.3.3). These were simple searches in which users selected one listed subject heading, displayed one or more titles, and gave "very useful" ratings to one or more titles. Only one of 44 successful keyword-in-subdivided-heading searches featured more than a little browsing.

Unsuccessful keyword-in-subdivided-heading searches outnumbered successful keyword-in-subdivided-heading searches in the Pinstripe System (section 11.3.4). Frequent reasons for failure were perseverance, database failure, and searches that retrieved no subject headings.

13.9.2 Free-text Approaches

Title-keyword searches were unique to the Blue System (section 11.4). Successful title-keyword searches outnumbered unsuccessful title keyword searches. Almost 70% of title keyword searches in UM-D's Blue System were successful; over 85% of title keyword searches in Earlham's Blue System were successful. On occasion, users rated several titles "very useful" that contained the same main or subdivided subject heading. A system that features relevance feedback could automatically examine subject headings

in useful titles. When title-keyword searches retrieve few titles, users could find additional material on their topics of interest from the system's retrieval of titles bearing these subject headings.

Both Blue and Pinstripe Systems featured keyword-in-record searches but the majority (95.3%) of keyword-in-record searches were conducted in the Pinstripe System (section 11.6). About half of keyword-in-record searches were successful. Keyword-in-record searches were low posted in the Blue System where they averaged 2.3 titles per search. Keyword-in-record searches were much higher posted in the Pinstripe System where they averaged 81.8 titles per search. Large retrievals were the major problem with keyword-in-record searches in the Pinstripe System. Often users walked away without rating displayed titles or terminated their searches after rating the first of many retrieved titles.

13.9.3 Subject Searching for Personal Names

No subject queries for personal names that users entered into the experimental online catalog at UM-D were usable names (section 11.7.2.1). Most were out-of-scope or playing around. Users also entered subject queries for personal names that were deemed unusable into the catalog at Earlham (section 11.7.2.2); however, a total of 27 usable queries were entered. All but four of these queries contained topical elements. Blue-system search trees for handling personal-name queries were especially effective in retrieving useful titles for user queries (section 11.7.2.4).

13.9.4 Spelling and Vocabulary Problems

The analysis of spelling errors in user queries demonstrated that online catalogs should feature computer-assisted spelling correction routines — just like word processing programs — to assist end users in identifying and correcting their spelling errors (section 12.2). Solving vocabulary problems (section 12.3) would require several enhancements to search tree configurations: (1) eliminating punctuation in queries including double posting hyphenated words in subject headings and title fields, (2) creating a go/see list to treat words and phrases as equivalent terms during retrieval, (3) automatic stemming of queries to effect retrievals in exact and the various keyword searches, and (4) best-match approach to retrieve titles for the most difficult user queries that traditional searches — exact, alphabetical, and keyword — cannot handle.

13.10 Redesigning Search Trees for Responsive Subject Searching

13.10.1 Redesigning Search Trees

Statistical and failure analyses resulted in many recommendations to improve the operation of specific subject searching approaches. They also resulted in recommendations to improve the design of search trees.

The only change to the initial search tree from the original design is to return users whose searches failed to produce retrievals to the question about personal names (see figure 12.3).

The search tree for the exact approach remains unchanged. Systems respond to user queries for subjects generally that match exact or normalized forms of controlled vocabulary terms with the exact approach (figures 12.4A and 12.4B).

The redesigned search tree for one-word queries contains a few changes from the original search tree for one-word queries: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names, (3) invoking the keyword-in-record search, (4) invoking relevance feedback following title-keyword and keyword-in-record searches, and (5) invoking the alphabetical approach as the search type "of last resort." The search tree for one-word queries is given in figures 12.5A and 12.5B.

The redesigned search tree for multi-word queries features the following enhancements: (1) checking go/see lists, (2) redirecting misspelled queries to the question on personal names, (3) performing relevance feedback based on useful retrievals in free-text searches (i.e., title-keyword, keyword in subject heading fields, and keyword-in-record searches), (4) invoking stemming, (5) invoking the best-match approach, and (6) invoking the alphabetical approach as the search type "of last resort." The redesigned tree is shown in figures 12.6A–C.

The search tree for personal-name subject queries remains unchanged from the original conception (Table 4.1). The first step is for systems to single out queries containing topics. These queries are submitted to the keyword-in-subdivided-heading search followed by the keyword-in-record search in the hopes of finding headings and bibliographic records containing both name and topic elements. If systems fail to find both elements, they omit the topic element from the query and continue searching through the alphabetical approach. Personal-name queries consisting exclusively of name elements are submitted to the alphabetical approach only.

13.10.2 Abridged Search Trees

Users who search small databases that feature search trees might find expanding their search somewhat tiresome. Thus, we recommend abridged search trees that do not feature as many subject searches and expanding options (sections 12.5.3 and 12.5.5). Abridged search trees for one-word queries feature only three subject searching approaches: (1) exact, (2) alphabetical, and (3) keyword-in-record searches with relevance feedback. The search tree for multi-word queries features five subject searching approaches: (1) exact, (2) alphabetical, (3) keyword-in-heading, (4) keyword-in-record with relevance feedback, and (5) best-match with relevance feedback.

13.11 A New Design for Subject Access to Online Catalogs

This project tested a new design for subject access to online catalogs. It compared the performance of an experimental online catalog that was governed in its selection of subject searching approaches by search trees (Blue System) with the performance of an experimental online catalog that chose subject searching approaches randomly (Pinstripe System). The random selection of subject searching approaches was intended to reflect how end users select subject searching approaches when systems provide them with multiple choices.

The Blue System was more likely than the Pinstripe System to provide opportunities for guiding users to useful information on their topics of interest than the Pinstripe System provided. In full administrations of the Comparison Search Experiment at UM-D, precision of Blue-system searches was significantly greater than precision of Pinstripe-system searches. In both Blue and Pinstripe Systems at UM-D and Earlham, controlled vocabulary searches fared significantly better than free-text searches. About three-quarters of Blue-system searches were successful in retrieving at least one title that users judged “very useful.” This definition of successful searches included sizable numbers of Blue-system searches that were marred by user perseverance and navigation problems. Recommendations were made to improve the design of the exact approach and minimize problems connected with perseverance and navigation.

End users rated the performance of the Blue System much higher than the Pinstripe System in terms of retrieving greater numbers of useful titles and making them satisfied with their search, and they preferred the Blue System over the Pinstripe System.

Overall, between 50% and 60% of Pinstripe-system searches were successful. Success in the Pinstripe System was negatively affected by the system’s random selection algorithm for subject searching approaches. Between ten and twenty percentage points separated

the results of Blue- and Pinstripe-system searches for search approach failures. The keyword-in-subdivided-heading search was usually to blame for search approach failures; however, the alphabetical approach had large shares of the failures. Subject searching approach failures were rare in the Blue System, accounting for only 1.3% of Blue-system searches. Search-approach failures accounted for the failure of 14.3% of Pinstripe-system searches. Pinstripe-system searches were also negatively affected by large retrievals — a problem that seldom occurred in Blue-system searches.

The analysis of Comparison Search Experiment data demonstrated that the search trees were more effective in selecting a subject searching approach that would produce useful information for the subjects users seek than users would select on their own. This project provided recommendations for improving specific subject searching approaches to increase their efficiency, increase user perseverance, and encourage browsing. It also suggested enhancements to the search trees to enable them to respond to the wide variety of user queries for subjects.

References

Drabenstott, Karen M., and Diane Vizine-Goetz. 1994. *Using subject headings for online retrieval: Theory, practice, and potential*. San Diego: Academic Press.

Mitev, Nathalie, Gillian Venner, and Stephen Walker. 1985. *Designing an online public access catalog*. London: British Library. Library and Information Research Report 39.

Siegel, Elliott R. et al. 1983. "Research strategy and methods used to conduct a comparative evaluation of two prototype online catalog systems." In *National Online Meeting proceedings —1983*, 1983 April 12–14, New York, compiled by Martha E. Williams and Thomas H. Hogan, 503–11. Medford, N.J.: Learned Information.

Appendix A. Form Subdivisions

Topical subdivisions assigned to subfield code $1 for form subdivisions in bibliographic records of the University of Michigan-Dearborn:

Abbreviations
Abstracts
Amateurs' manuals
Anecdotes, facetiae, satire, etc.
Atlases
Bibliography
Bio-bibliography
Biography
Book lists
Case studies
Catalogs
Charts, diagrams, etc.
Chronology
Congresses
Diagrams
Dictionaries
Directories
Drawings
Early works to 1800
Encyclopedias
Examinations, questions, etc.
Exhibitions
Film catalogs
Films
Formulae, receipts, prescriptions
Guide-books
Handbooks, manuals, etc.
Illustrations
Indexes
Job descriptions
Juvenile literature
Laboratory manuals
Maps
Miscellanea
Observers' manuals
Outlines, syllabi, etc.
Periodicals
Photograph collections
Pictorial works
Popular works
Portraits
Problems, exercises, etc.
Programmed instruction
Registers
Study guides
Tables
Textbooks for foreign speakers
Union lists

Topical subdivisions assigned to subfield code $1 for form subdivisions in bibliographic records of Earlham College:

Abstracts
Addresses, sermons, etc.
Anecdotes
Archival resources
Audio-visual aids
Bibliography
Bio-bibliography
Biography
Caricatures and cartoons
Cartoons, satire, etc.
Case studies
Cases
Catalogs
Chronology

Congresses
Controversial literature
Correspondence
Curricula
Designs and plans
Diaries
Dictionaries
Dictionaries, indexes, etc.
Directories
Early accounts to 1600
Early works to 1800
Encyclopedias
Exhibitions
Facsimiles
Festschriften
Fiction
Field work
Gazetteers
Glossaries
Glossaries, vocabularies, etc.
Guide-books
Handbooks, manuals, etc.
Humor, caricatures, etc.
Indexes
Interpretive programs
Interviews
Juvenile literature
Library resources
Lists
Manuscripts
Maps
Maps, Early American
Miscellanea
Outlines, syllabi, etc.
Pamphlets
Passenger lists
Periodicals
Photograph collections
Pictorial works
Pictures, illustrations, etc.
Popular works
Portraits
Portraits, caricatures, etc.
Registers
Registers of births, etc.
Reviews
Sermons
Sources
Statistics
Statistics, Vital
Tables
Textbooks
Texts
Vocabularies, etc.
Voting registers

Appendix B. Control-field Elements

Table B1. Other Contents: Illustration Code (USMARC Tag 008, Positions 18-21)

Code	English-language equivalent term
a	Illustrations
b	Maps
c	Portraits
d	Charts
e	Plans
f	Plates
g	Music
h	Facsimiles
i	Coats of arms
j	Genealogical tables
k	Forms
m	Phonodisc, phonowire, etc.
o	Photographs
p	Illuminations

Table B2. Book Details: Nature of Contents Code (USMARC Tag 008, Positions 24-27)

Code	English-language equivalent term
t	Technical reports
b	Bibliographies
k	Discographies
q	Filmographies
c	Catalogs
i	Indexes

a	Abstracts/summaries
d	Dictionaries
e	Encyclopedias
r	Directories
s	Statistics
f	Handbooks
p	Programmed texts
w	Law reports and digests
g	Legal articles
o	Reviews
v	Legal cases and case notes
n	Literature surveys

Table B3. Book Details: Festschrift Code (USMARC Tag 008, Position 30)

Code	English-language equivalent term
1	Festschrift

Table B4. Book Details: Biography Code (USMARC Tag 008, Position 34)

Code	English-language equivalent term
1	Contains biographical information

Appendix C.
End-user Questionnaire

The pre-search questionnaire consists of questions 1–3. The post-search questionnaire consists of questions 4–14. Preceding questions 1, 4, 7, and 13 are comments that introduce several questions.

Please begin by answering three questions about yourself:

1. How many times have you tested this system?
   - ( ) Never before
   - ( ) One time
   - ( ) Two times
   - ( ) Three or more times

2. How often do you use other computer systems?
   - ( ) Daily
   - ( ) About every other day
   - ( ) Weekly
   - ( ) Monthly
   - ( ) About four times a year
   - ( ) About once a year
   - ( ) Never

3. What is your major field of study?
   - ( ) Computer science
   - ( ) Engineering
   - ( ) Business and management
   - ( ) Education
   - ( ) History
   - ( ) Mathematics
   - ( ) Other humanities fields

( ) Other sciences fields

( ) Other social sciences fields

( ) None of the above

Please answer some general questions about the titles the Blue and Pinstripe Test Systems found for the subjects you typed in.

4. Comparing the Blue Test System with the Pinstripe Test System, did you find:

   ( ) A greater amount of useful titles in the Blue Test System

   ( ) About the same amount of useful titles in both systems

   ( ) A lesser amount of useful titles in the Blue Test System

   ( ) No useful titles in either system

5. In relation to what you were looking for, would you say your search was more satisfactory in:

   ( ) The Blue Test System

   ( ) The Pinstripe Test Systems

   ( ) No difference

6. The next time you need to conduct a computer search, will you want to use:

   ( ) The Blue Test System

   ( ) The Pinstripe Test Systems

   ( ) No difference

Please answer the next six specific questions about using the Blue and Pinstripe Test Systems.

7. Which system do you find easier to give instructions to, for example, to type in subjects, continue searches, make selections from lists?

   ( ) The Blue Test System

   ( ) The Pinstripe Test Systems

   ( ) No difference

8. Which system gives you more useful instructions in response to the subjects you type in?

   ( ) The Blue Test System

   ( ) The Pinstripe Test Systems

   ( ) No difference

9. Which system tells you what to do more clearly?

   ( ) The Blue Test System

( ) The Pinstripe Test Systems

( ) No difference

10. Which system is more helpful for giving you new ideas for searching for subjects?

( ) The Blue Test System

( ) The Pinstripe Test Systems

( ) No difference

11. Which system is more efficient in guiding you to useful information?

( ) The Blue Test System

( ) The Pinstripe Test Systems

( ) No difference

12. On a scale of 1 to 5 in which 1 means "not at all important" and 5 means "very important," rate the following features as to their importance in helping you search library computer systems like ASTUTE:

( ) Ease of giving instructions to the system

( ) Usefulness of the system's instructions

( ) New ideas that the system gives you for searching for subjects

( ) The system's clarity in telling you what to do

( ) The system's efficiency in guiding you to useful information

( ) The system's ability to continue searching on its own to find useful information

Please answer two final questions on the subjects you search in library computer systems.

13. How would you rate your familiarity with the subject you searched today?

( ) Very familiar

( ) Somewhat familiar

( ) Somewhat unfamiliar

( ) Very unfamiliar

14. How interested are you in seeing the capabilities of ASTUTE extended to topics besides [Engineering and Computer Science (at UM-Dearborn)] or [American History (at Earlham)]?

( ) Very interested [prompt with question 15]

( ) Somewhat interested

( ) Somewhat uninterested

( ) Very uninterested

15. Please type in what topics you would like added to ASTUTE.

Appendix D.
Cover Letter and Consent Form Completed by Library Staff

[Date]

[Name and address of
prospective library staff participant]

Dear [prospective library staff participant]:

You are invited to participate in a study to test a new subject access design to online catalogs. This study involves searching the ASTUTE online catalog that has been installed in the reference area of [Mardigian or Lilly] Library since [date] and answering questions about the search that ASTUTE and I ask you. If you wish to participate, please fill out and sign the enclosed form, and return it to me in the enclosed envelope by [date]. At the same time as you are completing the form, [name of library liaison] will contact you about setting up a one-hour appointment to search ASTUTE in [date]. ([Name of library liaison] is helping me schedule the test to save me from making several than trips to [Mardigian or Lilly] Library.)

Please bring three to five subject queries in the area of [computer science or engineering; American history] to search. These queries could have come from library patrons in the course of your reference work or may be subject queries in which you have a professional or personal interest. I will not reveal your identity in reports generated by this study but I may discuss your queries and ASTUTE's ability to retrieve relevant titles for them.

Thank you for your assistance. If you have any questions about the study, please feel free to contact me by phone at 313-763-3581 or by electronic mail (karen.drabenstott@umich.edu).

Sincerely,

Karen M. Drabenstott
Associate Professor

Enclosure (1)

SCHOOL OF INFORMATION AND LIBRARY STUDIES
NEW SUBJECT ACCESS DESIGN STUDY
Consent Form

Please print:

Name:

Title:

Mailing address:

Office phone:

Electronic mail address:

Do you prefer to be contacted by
electronic mail? (Circle one:) Yes No Don't care

Please read the following and sign:

I have read the enclosed information (cover letter) about this study and agree to participate.

I understand that I will not be identified in reports generated from this study. I understand that my query and ASTUTE's ability to retrieve relevant titles for it may be discussed in such reports.

I understand that my participation is voluntary, and that I may withdraw at any time.

Signed _____

Date _____

Please return this form to:

Karen M. Drabenstott
University of Michigan
School of Information and Library Studies
304 West Engineering Building
550 East University Avenue
Ann Arbor, MI 48109–1092

Appendix E. Post-search Questionnaire Administered to Library Staff

Questionnaire: # _____
Library: _ UM-D EC __ First system: _Blue Pinstripe
Date: _____ Time: _____ Interviewer: KD

(Note that more space was left in the actual interview form for participants' answers to open-ended questions to enable the interviewer to write down participant responses.)

A. Introduction

Hello. I'm _____ from the School of Information and Library Studies at the University of Michigan. We are conducting a study to find out how well the ASTUTE experimental online catalog helps library patrons and staff find library material on the topics that interest them. ASTUTE is actually 2 different online catalogs. The 2 ASTUTEs contain the same titles but they have different subject searching capabilities. As you search, take note of the differences between the two ASTUTEs so that we can talk about these differences later.

We'll take 1 topic at a time. When you start searching ASTUTE, the system will ask you some questions about yourself and your previous use of the system. Then it will ask you to enter your subject query in 1 of the 2 ASTUTE online catalogs. As you retrieve titles, ASTUTE will ask you to assess their relevance. When you are done searching and make some moves to start over, ASTUTE will switch to the second ASTUTE catalog and search for the original subject query you entered. After you finish searching, ASTUTE will ask you some questions about your searches and search results. Then I will ask you additional questions about your experiences searching the 2 ASTUTEs.

Let's get started. Please go ahead and start searching ASTUTE for 1 of the 3 topics you brought with you.

Time: _____

(Respondent searches for 8–13 minutes)

B. Questions for Search Topic 1.

1. What was the impetus for the subject you searched in ASTUTE?

(Probe: troublesome patron query, developing a bibliography, responsibility for collection development, frequent patron request)

2. Did you have any difficulties searching for your subject in the first (Blue or Pinstripe) system?

(Probe: Give examples.)

3. What improvements could be made to the first (Blue or Pinstripe) system?

4. Did you have any difficulties searching for your subject in the second (Blue or Pinstripe) system?

(Probe: Give examples.)

5. What improvements could be made to the second (Blue or Pinstripe) system?

6. What differences did you notice between the ways in which the 2 systems responded to the subjects you typed in?

a. Probe for examples.

b. Probe: What do you think about this?

c. Probe: Can you tell me why?

d. What bearing does this have on your system preference?

7. In what ways did the (Blue or Pinstripe) system give you new ideas for searching for subjects?

a. Probe for examples.

b. Probe: What do you think about this?

c. Probe: Can you tell me why?

d. What bearing does this have on your system preference?

8. What differences did you notice between the titles the 2 (Blue and Pinstripe) systems found for your subject?

a. Probe for amount of titles, proportion of relevant titles, proportion of nonrelevant titles.

b. What bearing does this have on your system preference?

Let's move onto the next search. Please go ahead and start searching the (Blue or Pinstripe) system for the second of the 3 topics you brought with you.

Time: _____
(Respondent searches for 8-13 minutes)

C. Questions for Search Topic 2.

(Repeat questions B.1–9 above.)

Let's move onto the last search. Please go ahead and start searching the (Blue or Pinstripe) system for the third topic you brought with you.

Time: _____
(Respondent searches for 8–13 minutes)

D. Questions for Search Topic 3.

(Repeat questions B.1–9 above.)

E. Summary Questions

Imagine that we could transform your library's online catalog into a new catalog using the best features of this experimental online catalog.

10. What do you like most about the Blue or Pinstripe system?

(Probe: which system, give examples of features.)

11. What do you like least about the Blue or Pinstripe systems?

(Probe: which system, give examples of features, feeling of control, etc.)

12. Overall, which system do you prefer?

a. Can you tell me why?

13. Are there any additional comments you would like to make concerning your use of ASTUTE?

THANK YOU for participating in this study.
Your comments will help us make recommendations
for improving subject access
in online catalogs.

Appendix F.
ASTUTE Transaction Log Records

F.1Level 1 Transaction Log Records

ASTUTE assigned participants unique identification numbers in the order in which they used the system. Level 1 transaction log records contained these unique identification numbers, start and end times for the user's interaction with ASTUTE, and answers to the pre- and post-search questionnaires. User identification numbers were the same across all four levels of transaction log records, and, thus, served as a link between records at different levels. ASTUTE wrote one level 1 transaction log record per system user. When users walked away from their search before finishing, the system stopped writing values to level 1 record fields and left them blank. Table F1 lists fields of level 1 transaction log records.

Table F1. Level 1 Transaction Log Record Fields

Field description	Field abbreviation	Field type	Field length
User identification number	U_Id_No	numeric	6
Date	L1_Date	character	8
Search experiment start time	L1_St_Time	character	8
Search experiment end time	L1_En_Time	character	8
First system searched	L1_Sys_Ord	character	1
Answer to pre-search question 1	L1_Q1	character	1
Answer to pre-search question 2	L1_Q2	character	1
Answer to pre-search question 3	L1_Q3	character	2
Time immediately prior to user's answer to question 4	L1_Q4_Time	character	8
Answer to post-search question 4	L1_Q4	character	1
Answer to post-search question 5	L1_Q5	character	1
Answer to post-search question 6	L1_Q6	character	1

Answer to post-search question 7	L1_Q7	character	1
Answer to post-search question 8	L1_Q8	character	1
Answer to post-search question 9	L1_Q9	character	1
Answer to post-search question 10	L1_Q10	character	1
Answer to post-search question 11	L1_Q11	character	1
Answer to post-search question 12, first-listed category	L1_Q12_1	character	1
Answer to post-search question 12, second-listed category	L1_Q12_2	character	1
Answer to post-search question 12, third-listed category	L1_Q12_3	character	1
Answer to post-search question 12, fourth-listed category	L1_Q12_4	character	1
Answer to post-search question 12, fifth-listed category	L1_Q12_5	character	1
Answer to post-search question 12, sixth-listed category	L1_Q12_6	character	1
Answer to post-search question 13	L1_Q13	character	1
Answer to post-search question 14	L1_Q14	character	1

Table F2 shows a level 1 transaction log record for ASTUTE user #6 at Earlham College. Answers to pre-search questions 1–3 tell us this user is using ASTUTE for the first time, uses other computer systems on a weekly basis, and has a major field of study in the humanities.

Table F2. Complete Level 1 Transaction Log Record

Field abbreviation	Answer code	Answer category
U_Id_No	6	(User identification number)
L1_Date	2/23/93	(Date)
L1_St_Time	21:02:32	(Search experiment start time)
L1_En_Time	21:13:37	(Search experiment end time)
L1_Sys_Ord	P	Pinstripe (first system searched)
L1_Q1	1	"Never before"
L1_Q2	3	"Weekly"
L1_Q3	7	"Other humanities fields"
L1_Q4_Time	21:11:30	Time immediately prior to user's answer to question 4
L1_Q4	2	"About the same number of useful titles in both systems"
L1_Q5	3	"No difference"

L1_Q6	3	"No difference"
L1_Q7	3	"Both are easy"
L1_Q8	3	"Both have useful instructions"
L1_Q9	3	"Both are clear"
L1_Q10	3	"Both are helpful"
L1_Q11	4	"Neither is efficient"
L1_Q12_1	5	5 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_2	4	4 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_3	4	4 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_4	5	5 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_5	5	5 on a scale of 1, "not at all important," to 5, "very important"
L1_Q12_6	4	4 on a scale of 1, "not at all important," to 5, "very important"
L1_Q13	4	"Very unfamiliar"
L1_Q14	2	"Somewhat interested"

Post-search questionnaire responses indicate that this user felt that there was no difference between the performances of the Blue and Pinstripe Systems. The user felt the systems were easy to use and understand but neither was especially efficient about retrieving useful information. Although he rated all capabilities listed in question 12 as being important, he felt that the ease of giving instructions to the system, system clarity, and system efficiency in guiding him to useful information were the most important aspects that helped him search library computer systems. The user was very unfamiliar with the subject he searched in the Blue and Pinstripe Systems and was somewhat interested in seeing the systems' capabilities extended to topics besides American history.

F.2Level 2 Transaction Log Records

Level 2 transaction log records documented the queries users entered into ASTUTE and the subject searching approaches the system enlisted to retrieve subject headings or titles. ASTUTE wrote as many level 2 records as necessary to document user queries and subject searching approaches. The user identification number field of level 2 records linked them to level 1, 3, and 4 records. The series number field of level 2 records linked level 2 records to one another. Table F3 lists fields in level 2 transaction log records.

Table F3. Level 2 Transaction Log Record Fields

Field Description	Field abbreviation	Field type	Field length
User identification number	U_Id_No	numeric	6
Level 2 series number	L2_Series	numeric	3
System user is searching	L2_Sys	character	1
Search approach in use	L2_M_Approach	character	1
User-induced approach	L2_Appr	character	5
Time level 2 record is written	L2_Time	character	8
Match made or not made using this approach	L2_Results	numeric	1
Number of titles retrieved	L2_Titles	character	5
System- or user-induced method of continuing the search	L2_Contype	character	4
User-entered query	L2_Query	memo	N/A

Level 2 records contained ten fields. User identification and series number fields linked level 2 records to each other and transaction records for other levels. Level 2 records documented the system the user is searching with a one-character code, i.e., B (Blue System) or P (Pinstripe System). The L2_Query field stored user-entered queries or user-selected terms in a variable-length memo field that contained as many characters as the queries or terms. The system wrote a one-character code in the L2_M_Approach field to document the subject searching approach the Blue or Pinstripe System used to produce retrievals for the user query. Approaches and codes are given in table F4.

Table F4. Codes Used in L2_M_Approach Fields

Codes	Subject search approach
1	Exact search
2	Alphabetical search
3	Keyword-in-main-heading search
4	Keyword-in-subdivided-heading search
5	Title-keyword search for queries exceeding one word
6	Keyword in subject heading fields search
7	Keyword-in-record search
8	Title-keyword search for one-word queries
9	Keyword-in-record search for personal names
A	Alphabetical search for personal names

0	(System switch to Blue or Pinstripe)

The system wrote a 1 (one) for success or 0 (zero) for failure to the L2_Results field to document whether the search approach coded in the L2_M_Approach field was successful. If the system was successful, it gave users options to further the search. For example, in response to a successful exact search, the system could give users several methods of furthering the search, e.g., displaying broader terms, displaying narrower terms, displaying scope notes, displaying subdivided forms of subject headings. The system wrote a code in the L2_Appr field to document the method that users selected to further the search. This field could also contain codes to indicate the type of subject search that the system was executing or the system's response to a user action to start over. The latter ran the gamut of subject searching capabilities implemented in the experimental online catalog. The former were limited to codes for "swit" (i.e., switch) to indicate that the Blue System was switching to the Pinstripe System or vice versa and to "quit" to indicate that the system was initiating the post-search questionnaire. Table F5 lists the three types of codes written to the L2_Appr field of level 2 records.

Table F5. Codes Used in L2_Appr Fields

Codes	Code Descriptions
Subject search codes:	
ex	System is executing the exact approach
alph	System is executing the alphabetical approach
kwmh	System is executing the keyword-in-main-heading search
kwt	System is executing the title-keyword search
kwsf	System is executing the keyword in subject heading fields search
kwr	System is executing the keyword-in-record search
pkwr	System is executing the keyword-in-record search for personal-name subject queries
palp	System is executing the alphabetical approach for personal-name subject queries
Codes for furthering searches:	
xref	Within the exact or alphabetical approach, the system is displaying a *see* reference
notes	Within the exact approach, the system is displaying a scope note
titles	System is displaying retrieved bibliographic records
nt	Within the exact approach, the system is displaying narrower terms
bt	Within the exact approach, the system is displaying broader terms
subj	Within the exact approach, the system is displaying subject headings subdivided by topical subdivisions

geog	Within the exact approach, the system is displaying subject headings subdivided by geographical subdivisions
chron	Within the exact approach, the system is displaying subject headings subdivided by chronological subdivisions
mat	Within the exact approach, the system is displaying subject headings subdivided by form subdivisions
Housekeeping codes:	
swit	System is switching between Blue and Pinstripe Systems
quit	System is initiating post-search questionnaire

The L2_Contype field documented the actions taken by users to continue the search. Table F6 lists and describes the codes the system wrote to this field.

Table F6. Codes Used in L2_Contype Fields

Codes	Subject search approach
next	Blue System automatically goes to next approach on search trees
more	User makes an action to further the approach. Examples: selecting headings, browsing narrower terms, displaying scope notes
brow	User browses a list but does not choose a listed subject heading
expd	User chooses to expand the search which induces the Blue System to go to next approach on search trees
prev	User goes to a previous screen in exact approach
swit	System switches to Blue or Pinstripe System
quit	User makes move to start over and system initiates post-search questionnaire
nfnd	System is unable to find user-entered query in keyword dictionary

Level 2 records also contained fields for the time of individual transactions. Numbers coded in the L2_Titles field designated the number of bibliographic records users retrieved for the queries they entered or subject headings they selected.

Table F7 shows the series of level 2 records linked to the level 1 record in Table F2. The user entered the query “richmond, indiana” into the Pinstripe System. The system wrote only one level 2 record for the search in the Pinstripe System. When the user made a move to start over, the system switched to the Blue System and executed a search for the original query, i.e., “richmond, indiana.” It wrote eleven level 2 records for user actions and system responses in the Blue System. In the level 2 records in Table F6, L2_Query fields bearing user-entered or user-selected terms are given under the nine other fields of these records.

Table F7. Complete Series of Level 2 Transaction Log Records

Id	Series	Sys	M_Ap	Appr	Cont	Time	Result	Titles
6	1	P	7	titles	next	21:03:05	1	16
richmond, indiana								
6	2	B	0	swit	swit	21:05:37	0	0
richmond, indiana								
6	3	B	1	ex	next	21:05:48	0	0
richmond, indiana								
6	4	B	2	alph	next	21:05:48	0	0
richmond, indiana								
6	5	B	3	kwmh	next	21:05:48	0	0
richmond, indiana								
6	6	B	4	kwsh	next	21:06:26	1	0
richmond, indiana								
6	7	B	4	kwsh	more	21:06:26	1	0
Historic buildings — Indiana — Richmond								
6	8	B	4	titles	more	21:06:40	1	1
Historic buildings — Indiana — Richmond								
6	9	B	4	kwsh	next	21:07:26	1	0
Historic buildings — Indiana — Richmond								
6	10	B	4	kwsh	expd	21:07:26	1	0
richmond, indiana								
6	11	B	5	titles	next	21:07:37	1	13
richmond, indiana								
6	12	B	5	kwt	quit	21:11:29	1	0
richmond, indiana								

Key to table headings:

Id = U_Id_No
Series = L2_Series
Sys = L2_System
Appr = L2_Approach
Cont = L2_Contype
Time = L2_Time

Result = L2_Result
M_Ap = L2_M_Approach
Titles = L2_Titles

Level 2 transaction log records demonstrate that the user searched the Pinstripe System first with the query "richmond, indiana." The system wrote code 7 in the L2_M_Approach field to document that the Pinstripe System randomly chose the keyword-in-record search for this query. The system retrieved sixteen bibliographic records (L2_Titles field). To determine whether the user displayed any of these records, we need to inspect level 3 records. The user made an action to terminate the search (L2_Contype field). In response, ASTUTE switched to the Blue Test System (L2_System field).

The Blue System executed the exact approach, alphabetical approach, and keyword-in-main-heading search (L2_Series, L2_M_Approach, and L2_Result fields) without success. The sixth record in the level 2 record series documented that the Blue System executed a keyword-in-subdivided-heading search (L2_M_Approach). The "1" in the L2_Result field indicates that the Blue System was successful in retrieving subject heading(s). The user selected a listed heading (L2_Contype). The heading is listed under coded data in the seventh record in the level 2 record series, i.e., "Historic buildings — Indiana — Richmond." This subject heading was connected with one title (L2_Titles field). The tenth record in the level 2 record series indicates that the user wanted the system to automatically expand the search (L2_Contype). This means the user pursued a system prompt that induced the Blue System proceed with the next approach on search trees to produce retrievals. The system resumed with a title keyword search (L2_M_Approach) and was successful in retrieving thirteen titles (L2_Titles field). The contents of level 3 records are needed to determine whether the user displayed one or more of the thirteen records. When the user made a move to terminate the search, the system stopped writing level 2 records and resumed with the post-search questionnaire. (Level 1 records documented questionnaire responses.)

F.3Level 3 Transaction Log Records

Level 3 transaction log records contained fields to document bibliographic records that users displayed and rated. They also contained fields to enable the researchers to link level 3 records to level 1 and 2 records. Table F8 lists fields in level 3 transaction log records.

Table F8. Level 3 Transaction Log Record Fields

Field Description	Field abbreviation	Field type	Field length
User identification number	U_Id_No	numeric	6
Level 2 series number	L2_Series	numeric	3
Level 3 series number	L3_Series	numeric	4
Time level 3 record is written	L3_Time	character	8
Bibliographic record control number	Bib_Series	numeric	5
User relevance assessment	L3_Relevance	numeric	1

Serial numbers of the bibliographic records that users displayed were given in the Bib_Series field. User relevance assessments were assigned to the L3_Relevance field. The meaning of codes 1–3 in L3_Relevance field was: (1) "not useful," (2) "possibly useful," and (3) "very useful." Of the fields bearing numbers linking level 3 records to corresponding level 1 and 2 records, unique user identification numbers were given in U_Id_No fields. Level 3 records contained level 2 series numbers (L2_Series) to enable the researchers to determine the particular user activity that led to the retrieval of bibliographic records. Time stamps were given in L3_Time fields.

Table F9 shows level 3 transaction log records written for a user at Earlham College. Level 1 and 2 records for this user were shown in Tables F2 and F7. Level 2 records demonstrate that the user retrieved bibliographic records in level 2 series records 1, 8, and 11 (Table F7). Thus, the user had the opportunity to display records as a result of searching activity connected with level 2 series records 1, 8, and 11.

Table F9. Complete Series of Level 3 Transaction Log Records

U_Id_No	L2_Series	L3_Series	L3_Time	Bib_Series	L3_Relevance
6	1	1	21:03:05	10883	2
6	1	2	21:03:20	10821	1
6	1	3	21:03:32	10819	1
6	1	4	21:03:41	10811	1
6	1	5	21:03:52	10868	1
6	1	6	21:04:02	10739	1
6	1	7	21:04:07	10747	2
6	1	8	21:04:20	10869	1
6	1	9	21:04:26	10804	2
6	1	10	21:04:38	10888	1
6	1	11	21:04:44	10508	1

6	1	12	21:04:51	10858	1
6	1	13	21:04:59	10871	2
6	1	14	21:05:07	10746	1
6	1	15	21:05:14	10864	1
6	1	16	21:05:29	10887	1
6	8	1	21:06:47	10821	1
6	11	1	21:07:38	10883	2
6	11	2	21:07:52	10821	1
6	11	3	21:07:56	10811	1
6	11	4	21:08:00	10868	1
6	11	5	21:08:11	10739	1
6	11	6	21:08:15	10747	2
6	11	7	21:08:25	10804	1
6	11	8	21:08:31	10888	1
6	11	9	21:08:35	10858	1
6	11	10	21:09:37	10871	2
6	11	11	21:09:43	10746	2
6	11	12	21:10:44	10864	1
6	11	13	21:10:56	10887	1

In the Comparison Search Experiment conducted by the user at Earlham College, the system wrote thirty level 3 records for displayed and rated bibliographic records. The user displayed all sixteen records that the Pinstripe System retrieved in a keyword-in-record search. The user took two minutes and twenty-four seconds to display and rate the usefulness of the sixteen records. He gave four records “possibly useful” ratings and twelve records “not useful” ratings. The user retrieved fourteen bibliographic records in the Blue System. The user gave a “not useful” rating to the one record retrieved through the keyword-in-subdivided-heading search (L2_Series field = 8). Of the thirteen records retrieved through the title-keyword search (L2_Series fields = 11), the user rated four records “possibly useful” and nine records “not useful.” It took the user three minutes and eighteen seconds to review these thirteen records. The fourteen records displayed in Blue-system searches were displayed in Pinstripe-system searches. User relevance assessments for two of the fourteen titles displayed in searches on both systems varied (Bib_series #10804 and #10746). (Chapter 8 described bibliographic records that users displayed most than once and the resolution of conflicting relevance assessments.)

F.4Level 4 Transaction Log Records

Level 4 records were written under two conditions: (1) users selected the "very interested" category in response to post-search question #14 on their interest in seeing ASTUTE's capabilities extended to other topics, and (2) users typed something into the pop-up box asking them for suggested topics. Level 4 records contained only two fields: (1) the unique user identification number to link level 4 records to level 1, 2, and 3 records (U_Id_No), and (2) the terms users typed into the pop-up box (L4_Topics). Table F10 lists fields in level 4 transaction log records.

Table F10. Level 4 Transaction Log Record Fields

Field Description	Field abbreviation	Field type	Field length
User identification number	U_Id_No	numeric	6
User-entered suggested topic(s)	L4_Topics	memo	N/A

User identification numbers were limited to fixed-length fields. User-entered suggested topics were saved in a variable-length memo field. The system did not write a level 4 record for the user at Earlham because he did not select the "very interested" category in response to post-search question #14. Several ASTUTE users at Earlham College did select this category and their suggested topics are given in Table F11.

Table F11. Examples of Level 4 Transaction Log Records

U_Id_No	L4_Topics
31	i don't know
35	religion/philosophy
92	2
102	Child abuse
113	AFRICAN AMERICAN WOMEN WRITERS IN HARLEM RENAISSANCE
146	life
151	Jump Rope Through the Ages
164	supreme court decisions
193	po;itics
201	women in reconstruction; civil war

Although subsequent chapters describe user-suggested topics, it is interesting to see the wide range of comments. User-suggested topics could be very broad (e.g., "life") or very specific (e.g., "Jump Rope Through the Ages"). They could also be unintelligible (e.g., "2"), contain typographical errors (e.g., "po;itics"), or say what was on the user's mind (e.g., "i don't know").