

Returning Classification to the Catalog

The concept of a classified catalog, or using classification as a form of subject access, has been almost forgotten by contemporary librarians. Recent developments indicate that this is changing as libraries seek to enhance the capabilities of their online catalogs. The Western North Carolina Library Network (WNCLN) has developed a "classified browse" feature for its shared online catalog that makes use of Library of Congress classification. While this feature is not expected to replace keyword searching, it offers both novice and experienced library users another way of identifying relevant materials.

Classification to modern librarians is almost exclusively a tool for organizing and arranging books (or other physical media) on shelves. The role of classification as a form of subject access to collections through the public catalog—the concept of the classified catalog—has been almost forgotten. From a review of the literature, it does not appear that any major U.S. library has supported a classified catalog since Boston University Libraries closed its classified catalog in 1973.¹ To be sure, nearly all online catalogs nowadays have some form of what is called a “call number search” or a “shelf list browsing capability” that is based on classification, but this is a humble and little-used feature because it requires that a call number (or at least a call number stem) be known and entered by the user, when no verbal index to the classification is available online. This search methodology provides

Robert N. Bland (bland@unca.edu) is Associate University Librarian for Technical Services, University of North Carolina at Asheville. **Mark Stoffan** (mstoffan@fsu.edu) is Associate Director for Library Technology at Florida State University, Tallahassee.

nothing in the way of a systematic and hierarchical arrangement and display of subject classes, complete with accompanying verbal descriptions, that the classified catalog seeks to accomplish. But as Karen Markey put it in her recent review of classification and the online catalog, “To this day, the only way in which most end users experience classification online is through their online catalog’s shelf list browsing capability.”²

There are signs that this situation is changing. The recently released Endeca-based catalog at North Carolina State University Libraries uses Library of Congress Classification (LCC) in a prominent way to provide for browsing of the collection without need of the user entering any search terms at all.³ The LCC outline is presented on the main search entry screen with verbal captions describing the classes, allowing users to navigate through several layers of the outline to retrieve with a click of the mouse bibliographic records for materials assigned to those classes. In a converse way, the new online catalog being developed by the Florida Center for Library Automation uses LC classification as

a kind of back end to keyword searching. Following a keyword search, a user can limit the results set by confining it to a designated LCC range chosen again from an online display of the LCC outline.⁴ Both of these catalogs use three levels of the LCC outlines from the most general single letter level classes (Q for sciences, for example) through the two-letter classes for more specific subjects (QC for physics, QD for chemistry) to an even finer granularity with designated numeric ranges within the two-letter classes identifying specific subdisciplines, (QD241–QD441 for organic chemistry).

The Western North Carolina Library Network (WNCLN) has been experimenting with classification as a retrieval tool in the public catalog for some time,⁵ and it has just implemented the first version of what we call a Classified Catalog Browse in our Innovative Millennium system.⁶ Like the two catalogs just mentioned, the Classified Catalog Browse is based on software that is external to the ILS software and integrated with that software through linking and webpage designs. Also, like the previously discussed catalogs, it is



Figure 1. Level 1 of LC Classification in WNCLN WebPac

based on scanning and incorporating into the catalog the LCC outlines as published by the Library of Congress. The WNCLN catalog goes a step further, however, in bringing the entire LC classification online down to the individual class number level—at least that portion of the classification that is actually used in our catalog. This is done through extracting class numbers and associated subject headings from bibliographic and authority records in our catalog and building an online classification display with descriptive captions (a verbal index) from these bibliographic and authority records. The result is a hierarchical display (to continue the example from above) not only of QC241–QD441 for organic chemistry but within this, QD271 for chromatographic analysis, QD273 for organic electrochemistry, and so on. The design of our interface presents this as a fourth level to which the user can “drill” down beginning with Q for sciences, QD for chemistry, QD241–QD441 for organic chemistry, and finally QD271 for chromatographic analysis (figures 1–4). From this fourth level, the user can click an associated link to execute a search of the catalog by the class number in question using the call number search function of the ILS (figure 5); a second link for that class number will present the same list of titles but sorted by “Most Popular” (i.e., the items that have been checked out most frequently) from a separate but linked external database (figure 6); a third link will search the catalog by the associated subject heading for the class (figure 7); and finally a fourth link will show other subject headings that have been used in the catalog with this specific class number (figure 8).

What does having the LC classification online in our catalog accomplish for our users? Part of the point of our project is to answer this very question. Chan and others⁷ have theorized that incorporation of the classification system into the catalog as a retrieval tool can

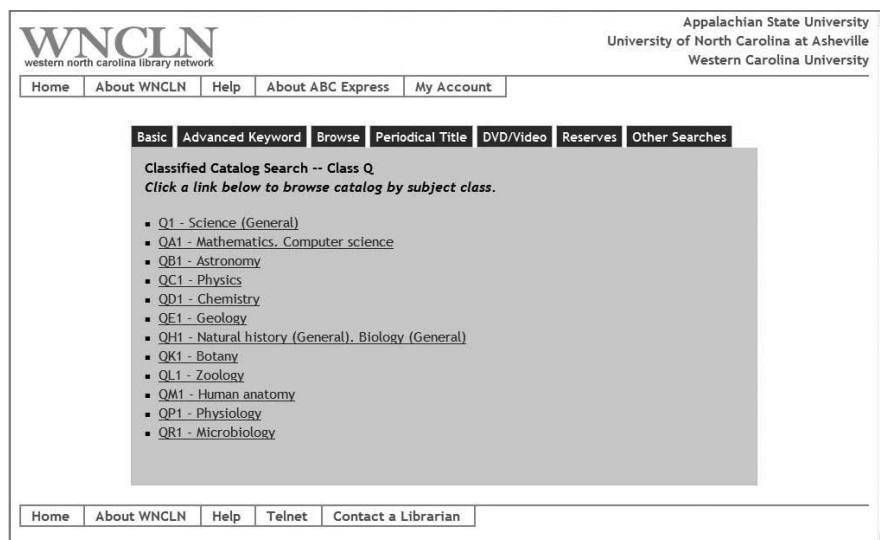


Figure 2. Level 2 of LC Classification in WNCLN WebPac



Figure 3. Level 3 of LC Classification in WNCLN WebPac

provide enhanced subject access that is not possible through standard alphabetical subject headings and keyword searching alone. Early studies by Markey and others at OCLC seem to have confirmed this with an online version of the Dewey Decimal Classification.⁸ Since (as far as we

know) the Library of Congress classification has not really been tested as an online retrieval tool in a live catalog up to now, our implementation will serve as a kind of test bed for this hypothesis. How actual users in fact exploit this feature is of course only something that experience will

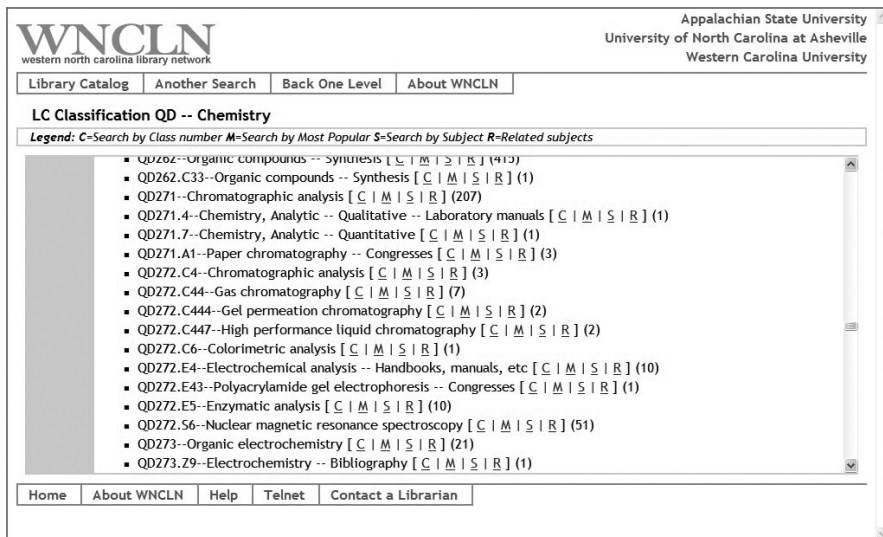


Figure 4. Level 4 of LC Classification in WNCLN Web:Pac

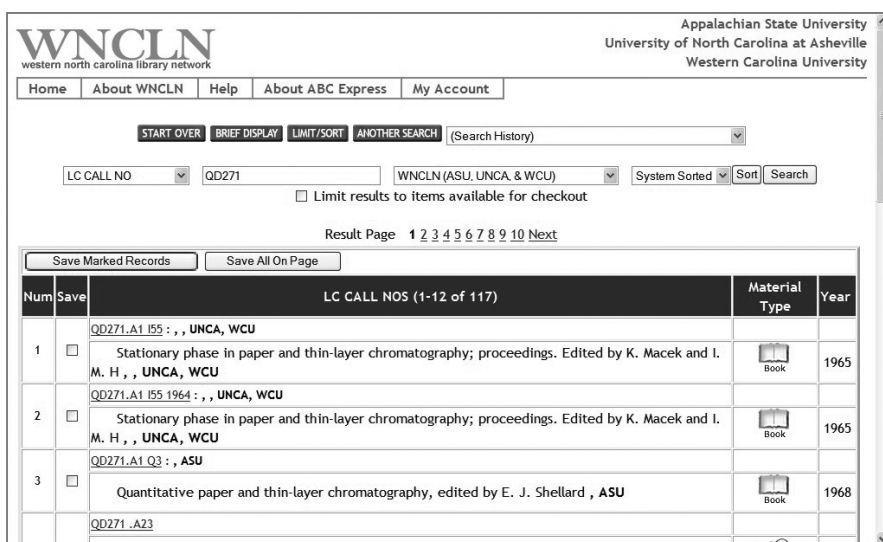


Figure 5. Call number search display in WNCLN

tell. A cursory look, however, would seem to indicate definite advantages to this approach.

First of all, many studies indicate that two of the major sources of failure with subject retrieval in online systems are misspellings and poor choice of search terms by users. No

matter how far we may try to go with keyword searching and relevance ranking, no online library retrieval system is likely to do much with "Napolyan's fites" when what the user is looking for are books on the military campaigns of the Emperor Napoleon. With the classification sys-

tem and verbal index online most of these problems are eliminated, since users can navigate to a subject of choice without ever entering a search term. Moreover, given the design of the verbal index based on Library of Congress subject headings, the user is led to actual subject headings used in the catalog, which should provide for precise retrieval beyond what is ordinarily possible with keywords even when entered correctly, and (importantly) a retrieval set that is always greater than zero. The infamous and frustrating problem of "no hits" is eliminated.

Secondly, the great attraction of the classified catalog approach is that it arranges subjects in a hierarchical fashion based on integral connections among the topics in a way that cannot be accommodated in an alphabetical subject approach because of the vagaries of spelling. The topics "Violence," "Social Conflict," and "Conflict Management," for example, obviously spread out in an alphabetical subject list, are collocated in the classified catalog under the class "HM1106-HM1171 Interpersonal Relations" (figure 9), allowing the user to find references to materials all in one place in the catalog just as the classification system arranges the books on these subjects all in one place on the library shelves. Alphabetical subject indexes, of course, attempt to ameliorate this problem by means of cross references, but there is clearly a limit to how far one can go with this approach.

Finally, the classified catalog provides an efficient way for collection development staff to review specific subject areas and to make better informed purchasing decisions regarding the collections. In the WNCLN design, the classes at the bottom level of the hierarchy are linked to the catalog by call number and subject headings, and each class carries an indication of the number of items assigned that class number. The classes are also linked to an external database that shows the frequency

of circulation of items in the class as well as title and date of publication. A quick review of this list can inform a bibliographer of circulation rates as well as the currency of materials in the class.

As mentioned, the captions that are displayed with the LCC hierarchy in the WNCLN catalog are extracted from subject headings and authority records present in our catalog. Readers familiar with LC MARC record services may wonder why we took this approach to building the verbal index rather than using the information available in the LC MARC classification records. Machine-readable records for LC classification are now available in MARC format. These files include records for each individual class number with a corresponding verbal caption. While we did experiment with using these files, cost and complexity determined that we go another direction. The LC classification files are huge, containing hundreds of thousands of classification numbers that we do not now and probably never would use in our WNCLN catalog simply because we (unlike LC) have no materials on these subjects. While these records could be filtered out by matching against LC class numbers that are found in our catalog and discarding non-matches, this would add yet another level of processing to an already complex process, as would handling the LC table subdivisions that are used in the LC schedules and that are separate from the standard class numbers. Secondly, the LC MARC classification files require a subscription costing several thousands dollars per year, as well as a substantial payment for the retrospective file needed to begin building the database of class numbers.

On the other hand, extracting the verbal index from subject headings and authority records in our own catalog adds no cost to our processing. These headings and authority records are created and maintained, of course, as a standard part of the

The screenshot shows the WNCLN website interface. At the top, it says 'WNCLN western north carolina library network' and 'Appalachian State University University of North Carolina at Asheville Western Carolina University'. There are navigation links: Home, About WNCLN, Help, About ABC Express, My Account. The main heading is 'Most Used Titles in Class QD271'. Below it is a table with columns: Rank, TITLE, Lib, Uses, Year.

Rank	TITLE	Lib	Uses	Year
1	Handbook of chemical property estimation methods : environmental behavior of o	UNCA	9	1982
2	A handbook of organic analysis : qualitative and quantitative / H. T. Clarke	WCU	7	1975
3	Paper & thin layer chromatography and electrophoresis a teaching level manual	WCU	5	1965
4	Handbook of chemical property estimation methods : environmental behavior of o	WCU	5	1982
5	Gas chromatography: principles, techniques, and applications [by] A. B. Little	WCU	5	1970
6	Interpretation of mass spectra of organic compounds [by] Mynard C. Hamming [an	UNCA	4	1972
7	Interpretation of mass spectra of organic compounds [by] Mynard C. Hamming [an	ASU	4	1972
8	Infrared spectroscopy [by] Robert T. Conley	WCU	4	1966
9	Gas chromatography. Edited by C. G. Verver. With a foreword by A. J. P. Martin	ASU	4	1959
10	Gas chromatography [by] L. Szepeszy [revised by P. Benedek translated from th	WCU	4	1970
11	Chromatographic methods, by R. Stock and C. B. F. Rice	WCU	4	1967
12	The characterization of organic compounds [by] Samuel M. McElvain	ASU	3	1953
13	Thin-layer chromatography. With a foreword by J. G. Kirchner	WCU	3	1963
14	The practice of gas chromatography, edited by Leslie S. Ettre [and] Albert Zla	WCU	3	1967
15	The mass spectra of organic molecules. [By] J. H. Beynon, R. A. Saunders and A	WCU	3	1968

Figure 6. Most used titles display

The screenshot shows the WNCLN website search interface. It includes the WNCLN logo and navigation links. The search bar contains 'Chromatographic analysis' and 'WNCLN (ASU, UNCA, & WCU)'. Below the search bar, there are options for 'System Sorted' and 'Sort'. The search results are displayed in a table with columns: Num, Save, SUBJECTS (1-12 of 18), Year, Entries (114 Found).

Num	Save	SUBJECTS (1-12 of 18)	Year	Entries (114 Found)
1		Chromatographic Analysis -- 9 Related Subjects		9
2	<input type="checkbox"/>	Chromatographic Analysis		68
3		Chromatographic Analysis Apparatus And Supplies -- See Chromatographic analysis Equipment and supplies		1
4	<input type="checkbox"/>	Chromatographic Analysis Bibliography : , WCU	1977	1
5	<input type="checkbox"/>	Chromatographic Analysis Congresses		4
6	<input type="checkbox"/>	Chromatographic Analysis Data Processing : , WCU	1983	1
7	<input type="checkbox"/>	Chromatographic Analysis Dictionaries		3
8	<input type="checkbox"/>	Chromatographic Analysis Encyclopedias : , , ASU, WCU	c2001	1
9	<input type="checkbox"/>	Chromatographic Analysis Equipment And Supplies Maintenance And Repair : , , ASU, WCU	1972	1
10		Chromatographic Analysis Equipment Industry -- See Also Chromatographic analysis Equipment and supplies		1

Figure 7. Subject search display in WNCLN

cataloging process, and accordingly only headings and authority records that match materials owned by our libraries are included. The description or caption that is finally assigned to a class number is determined by a computer program that analyzes both authority records and bibliographic records found in our catalog

that are assigned the class number in question, with the subject heading that is used most frequently as a primary subject generally being the one normally selected as the caption for the class. These class numbers with associated subject headings are processed then by another program, which eventually builds HTML files

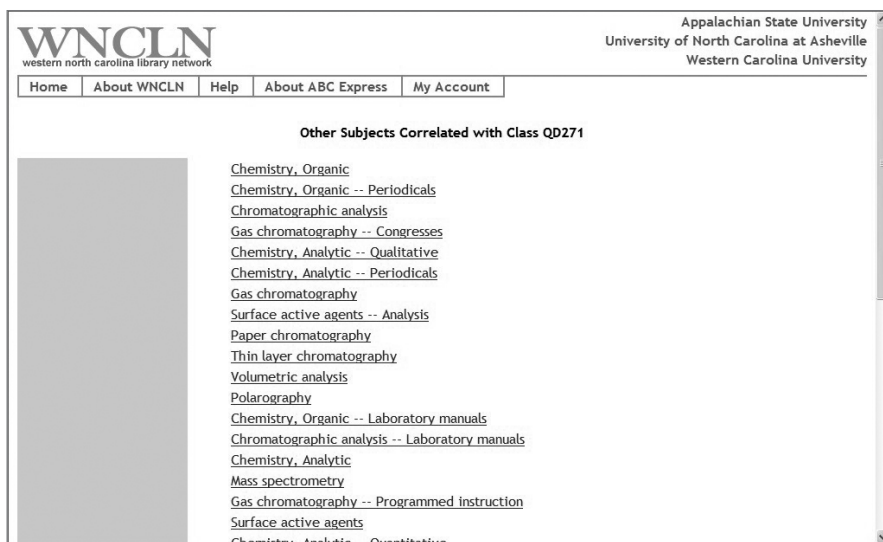


Figure 8. Related subjects display in WNCLN

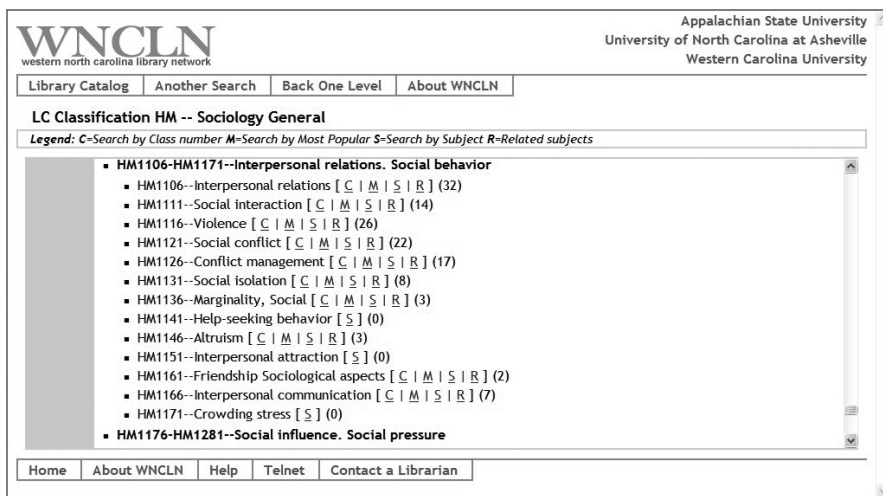


Figure 9. Collocation of terms in the classified catalog

representing the classification with links to the catalog and the external "Most Used" database as alluded to above. These standard HTML files, along with the files representing the first three levels of the LCC outline, are then loaded onto our Web server to display the classification system online.

A second advantage of this approach is that using the actual subject heading as the caption or description for the class makes it possible to use that caption as a direct link to a subject search in the catalog, as shown in the illustration in figure 4. A disadvantage is that the captions from the LCC files are designed to

retain the hierarchy that is represented in the printed schedules in a visual way by formatting and indenting. Captions derived from subject headings do not retain this feature. We have tried to accommodate this in our display of the schedules by replicating the class number ranges from the outline in the appropriate place in the full display of the schedules, thereby building a hierarchy from these ranges as genus and the individual class numbers as species. This does not manage to retain the full hierarchy of the LC schedules as shown in the printed schedules or as represented in LC's online Classification Web product, but it is, we hope, an adequate surrogate for the purpose intended. In fact, in most cases, the captions derived from the extracted subject and authority headings match quite nicely the captions included in the actual LCC schedules, as shown in a comparison from the psychology classification of the hierarchy as it appears in our Classified Catalog Browse and as it appears online in LC's Classification Web product (figures 10 and 11). What is missing in our representation of the classification is not so much the subject content of the classes but the notes and information about literary form that are included in the actual LCC schedules.

Thus, our LCC online is not a strict image of the LCC as it would appear in printed or electronic form based on the hierarchies and captions devised by the LC. Nor for that matter—despite our terminology—is it a true classified catalog, since only one classification (that used in the call number) is assigned to each item, whereas in a true classified catalog multiple classifications may be assigned to an item. It is nevertheless an online presentation of the LCC with links to our catalog that seeks to enhance subject access by exploiting the power of the classification system to organize materials by integral subject classes and to show relationships among subjects by a

hierarchical arrangement of classes as genus, species, and subspecies. And, perhaps just as importantly, it is an implementation that requires no additional cataloging effort on the part of our staff, nor any additional costs for data or processing other than the investment we have made in development of the software and the small amount of time required weekly to update the files.

We do not expect that the Classified Catalog Browse will replace keyword or subject searching as the primary means of subject access to our collections. We do believe that it promises to be a powerful and effective complement to our standard ILS searches that may improve subject searching for both the novice and the experienced user.

References

1. Margaret Hindle Hazen, "The Closing of the Classified Catalog at Boston University," *Library Resources and Technical Services* 18 (1974): 221-26.
2. Karen Markey, Joan S. Mitchell, and Diane Vizine-Goetz, "Forty Years of Classification Online: Final Chapter or Future Unlimited?" *Cataloging and Classification Quarterly* 42 (2006): 1-63.
3. North Carolina State University Libraries, "NCSU Libraries Online Catalog," North Carolina State University, www.lib.ncsu.edu/catalog (accessed Mar. 23, 2007).
4. Florida Center for Library Automation, "State University Libraries of Florida-Endeca," Board of Governors, State of Florida, <http://catalog.fcla.edu> (accessed Mar. 23, 2007).
5. The Western North Carolina Library Network is a consortium consisting of the libraries of Appalachian State University, the University of North Carolina at Asheville, and Western Carolina University.
6. Western North Carolina Library Network, "Library Catalog," Western North Carolina Library Network, <http://wncln.org> (accessed Mar. 23, 2007).

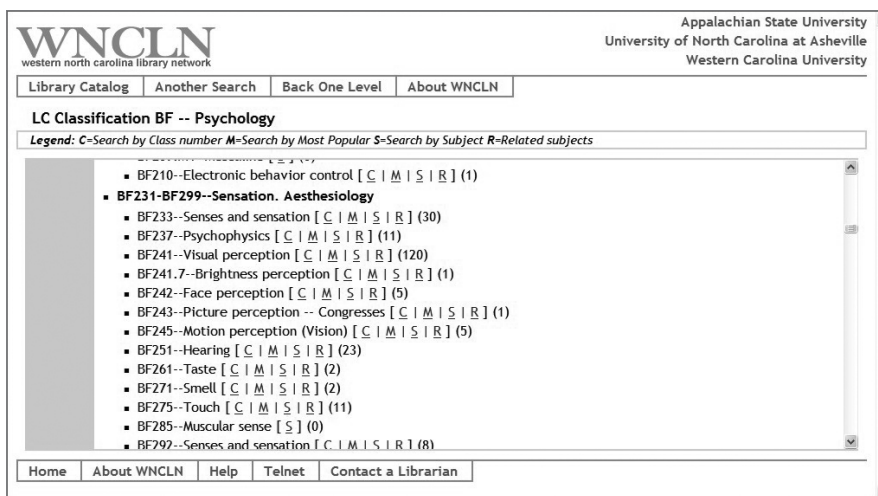


Figure 10. Class captions in the WNCLN WebPac

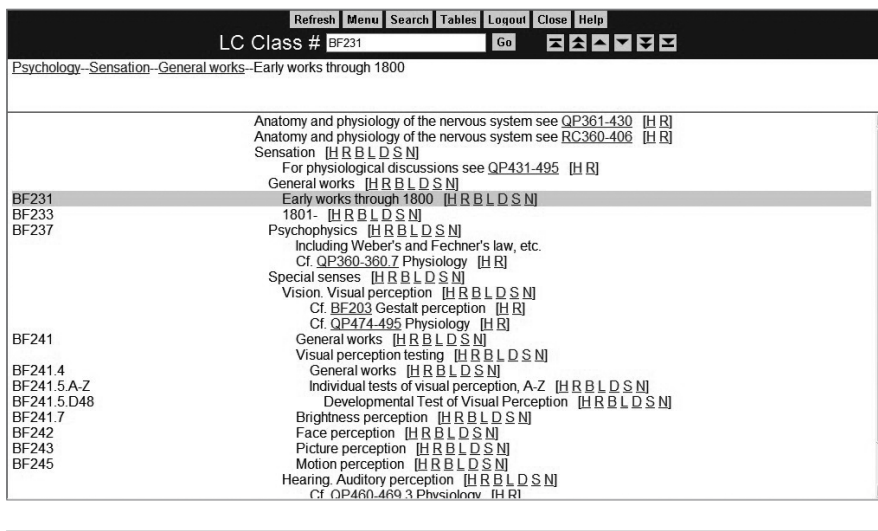


Figure 11. Class captions in LC's Classification Web

7. Lois Mai Chan, "Library of Congress Classification as an Online Retrieval Tool: Potentials and Limitations," *Information Technology and Libraries* 5 (1986): 181-92.
8. Karen Markey and Anh Demeyer, *Dewey Decimal Classification Online Project:*

Evaluation of a Library Schedule and Index Integrated into the Subject Searching Capabilities of an Online Catalog: Final Report to the Council on Library Resources (Dublin, Ohio: OCLC, 1986), Report no. OCLC/OPR/RR-86/1.