Factors in the Selection of a Classification Scheme for a Large General Library

This paper concerns the British Library; by now it is well known that the British Library consists of more than one large library. One of its components, the British Museum Library, is among the largest in the world; two others, the Science Reference Library in London, and the Lending Division in Yorkshire, both have collections sizable enough to raise problems regarding classification. In fact, however, the Lending Division has long since rejected classification as an operational tool in any other than peripheral uses, and is not a component which enters into the considerations I am making today.

In addition to these, there is the Bibliographical Services Division, which is evolving from the formerly separate entity, the British National Bibliography (BNB). This division is the most significant user of classification at the present time. Not only is the British National Bibliography primarily dependent on classification for its arrangement and articulation, but the classification data supplied by BNB is a substantial element of the centralized cataloging service which it gives to all kinds of libraries in Great Britain. One of the functions of the Bibliographical Services Division is to extend this service to cover the needs of in-house bibliographic processing within the British Library itself.
I am therefore dealing not so much with a large general library as with a large and complex national library system. Moreover, this is a system in which all the components have had a prior existence as independent organizations, sometimes with a long history of service and administration of their own. This aspect of the matter creates problems of rationalization which are perhaps unusual in terms of their scale, but which may not be dissimilar from those which arise wherever uniform bibliographic control is to be imposed on any two or more existing libraries which have long-established prior commitments and investments in their own different forms of control.

What makes the British Library a particularly interesting case for the classificationist, in my opinion, is that its principal prior commitment is the classification service which it provides for other libraries, through the British National Bibliography and its contributions to the MARC record. Contrast this with the Library of Congress, for example, where classification policies and developments have historically originated primarily to meet in-house requirements, and have only secondarily been determined by the needs of other libraries using the data. The British Library begins with a service to other libraries, and one of the main questions is whether that commitment can be extended to embrace its own, newly conceived in-house problems.

We know what consequences may spring from too close an adherence to the dictum that "what's good for General Motors is good for the country." Essentially, the British Library has to find an answer to the question: Can what is good for the country also be good for General Motors?

The Working Party

In 1972, a Working Party was established with the following terms of reference: "to examine the various classification and indexing systems currently in use in the various component parts of the British Library and to consider the possibilities of rationalisation, taking into account the need for standardisation nationally and internationally." The Working Party consisted of senior staff members with responsibilities for classification and indexing policies and for programs in each of the various component parts of the British Library (BL) together with two external members: Herbert Coblans—a distinguished authority on classification and indexing in the international sphere, and myself, who was honored with an invitation to act as chairman of the Working Party. The research department of Aslib acted as consultants on technical questions.

The Working Party delivered its final report to the British Library Board in June 1974; this is projected for publication in 1976, together with the texts of those supporting studies which the Working Party appended to the
report. In this paper I am therefore anticipating publication of the report to some extent, but my intention is to draw attention to those findings and conclusions of the report which seem to be of interest to classificationists and library administrators concerned with the problems of rationalization in this field.

I must begin this task by making the essential disclaimer that, although the authority for my comments derives from knowledge gained as chairman of the Working Party, nothing in what follows should be regarded as representing the official viewpoint or policy of the British Library Board. For one thing, the selection and interpretation of the Working Party's findings are my own. In addition, the British Library has not yet given a public indication of its endorsement of any of the recommendations made in the report.

The Problem

The only objective of rationalization is to achieve optimum cost-effectiveness. The ideal state of rationalization is one in which all requirements are met by a single system, generating the necessary data from a single source. The ultimate solution for the British Library would therefore be to find a single classification scheme which would yield maximum efficiency with respect to the following needs:

1. *The arrangement of books in the British Museum Library* The British Museum Library (BML) previously has used no classification for the arrangement of its stock. However, there are plans for it to move into a new, custom-designed building sometime in the late 1980s; when that happens, it will place on open access approximately one-quarter million volumes in the fields of humanities and social sciences. For these it needs a suitable classification. The great bulk of its vast collections will remain on closed access, but the availability of a class number for all acquisitions would enable it to exercise the greatest flexibility and economy in redetermining the contents of its open-access collections from time to time.

2. *The arrangement of books in the Science Reference Library* At present, virtually all Science Reference Library (SRL) stock is on open access, and classified in accordance with a special classification developed within the library itself from an earlier Patent Office Library classification. It is presently housed at two separate main sites, but it will eventually occupy one wing of the new British Library building and will thus exist physically next to the British Museum Library, with quick and easy access by readers from one library to the other.

3. *The arrangement of bibliographic records in the subject catalogs of both BML and SRL* At present, the principal catalogs are the published British
Museum Subject Index, which covers the fields of the British Museum Library, and at SRL, the card catalogs, arranged in accordance with SRL’s own classification scheme. It must be remembered, of course, that BML shares with SRL the task of conserving the British copyright deposit intake, along subject-divided lines, and there is an expectation that the published British Museum Subject Index may be extended in scope to embrace SRL’s work in its own fields of responsibility.

4. The arrangement of records in the British National Bibliography Arising largely from this need, and of equal weight in the Working Party’s terms of reference, are the two following requirements.

5. National standardization BNB and British MARC act as sources of centralized cataloging and classification data for a large number of academic and public libraries throughout the United Kingdom and elsewhere. The classification data used to arrange BNB, and provided in MARC, should meet the needs of shelf arrangement and bibliographic records in the greatest possible number of other libraries within the national network.

6. International Standards As the principal national library, the British Library is increasingly involved in the interchange of bibliographic information at the international level. Classification and other subject data are a significant aspect of this interchange, in regard to the cost-saving utilization by all exchange partners of the information flowing through the international networks.

Indexing and Information Retrieval

We began in the sphere of indexing and information retrieval. We examined comprehensively the European and North American literature reporting experimental work or summarizing the present state of the art on mechanized searching and retrieval by means of MARC tapes or by Dewey Decimal Classification (DDC), Library of Congress Classification (LCC), or Universal Decimal Classification (UDC). We also studied problems, such as profile construction, that are raised by the use of classification in these spheres. This is a large and difficult area, on which it would be possible to spend the remainder of this session. I will restrict myself to reporting our conclusions.

A satisfactory basis for a machine information retrieval system cannot be established without first identifying comprehensively the range of services which it is to supply, and then investigating the particular problems of each service. At the British Library, the potential demand for such services is very large and diffuse and, at this stage of development of BL’s internal and
external connections, the problems of coordination and integration seemed to us nearly insurmountable.

In addition, we felt that experimental work to date demonstrates that *verbal* mechanisms are superior to classificational notation mechanisms in achieving effective subject retrieval systems.

Undoubtedly this second conclusion was colored by the fact that PRECIS (the Preserved Context Indexing System) already incorporates a "verbal" machine-based indexing system that has been used for four years by the *British National Bibliography*, and recently by some library institutions, to create subject indexes for classified arrangements of document citations.

PRECIS was, in fact, the only available machine-based indexing system that appeared to have the potential to meet whatever future requirements might arise in this sphere in the British Library. We saw it first as a means of dealing with the immediate problem of the *British Museum Subject Index*, the production of which, in its present form, involves expensive manual elements and has been increasingly delayed because of staffing difficulties. A preliminary investigation showed that it was possible to manipulate PRECIS strings to produce mechanically an acceptable subject heading system which could replace the *British Museum Subject Index* with gains to the user, more efficient indexing information, and without an increase in cost. Further testing was undertaken and (subject to its satisfactory conclusion) we recommended that the PRECIS subject heading system be adopted to provide subject access to the BML collection and elsewhere as required. We saw it as the only means of providing a unified subject index to the BL Reference Division collections, and ultimately PRECIS strings would be added to all the records in the data base of the BL Reference Division.

**Classification and Interlibrary Relations**

Turning now to classification policies for the British Library in general, and for the BNB in particular, I have already emphasized that these cannot be determined without reference to the external use of classification schemes, nationally and internationally. The BL is committed to making its bibliographic data base available to other libraries and will itself receive large quantities of bibliographic records from other libraries for its internal use. There is, therefore, a potentially large demand from the library community for the provision of standard classification notations on British Library records. There may also be significant savings in the use of classification marks available on foreign records. The current ferment of activity throughout the library world in these areas suggested that further investigation of developments at a later date, when clearer pictures emerged, would be required before
a final decision was made. Looking at it now, eighteen months later, I do not think that the pictures have yet become any clearer.

For the picture as it appeared then, we collected and analyzed information on the comparative use of the published general classification schemes, in British libraries and in the national bibliographies of thirty-three countries. From our analysis, there is no question of the predominance of DDC in Great Britain. We established that DDC is used by 47 percent of all libraries,* and that it controls the arrangement of 75 percent of all library holdings. In contrast to this, UDC is used by 22 percent of British libraries, but controls the arrangement of only 5 percent of library holdings. As for LCC, particularly favored by British university libraries, it is used by 2 percent of all our libraries and controls the arrangement of 6 percent of library holdings.

Internationally, it appeared that DDC and UDC have each an equal number of users at the level of the national bibliography or national library agency. We estimated that the annual output of authoritative machine-readable records which carry DDC numbers was about 131,000; there was a similar quantity of machine-readable records with LCC numbers, but none carried UDC numbers. On this evidence, we made a firm recommendation that the bibliographic records created in BL's Bibliographical Services Division should continue to carry both DDC and LCC classmarks, and that they should do so as long as these facilitated the supply of exchangeable MARC records, and the generation of classified catalogs and bibliographies in forms acceptable and useful to public and academic libraries.

In recent years, there has been some lobbying in Great Britain for the addition of UDC numbers to the MARC data base, but the evidence we obtained of national usage did not support it very strongly. We concluded that UDC should only be added to DDC and LCC numbers if the British Library found it desirable to do so for its own purposes—that is, to facilitate information exchanges with other national libraries in Europe or elsewhere, or to provide a basis for its own shelf arrangement.

Classification for Shelf Arrangement

We now come to what proved to be the most difficult part of our brief: the determination of classification policies for the two great libraries of the Reference Division—the British Museum Library and the Science Reference

---

*It will be noticed that this figure differs appreciably from that reported by Down- ing on p. 73 of this volume. It appears that there is no one authoritative list of British libraries, and that the British Library survey reported by Lewis was done by Aslib and included all special libraries which were institutional members.—Ed.
Library. These two libraries have developed quite separately, on separate sites, with quite different histories, and with different operating policies.

The British Museum Library’s particular need was for a classification scheme appropriate for the open-access collection, covering all fields of humanities and social sciences. Since it has not been classified before, there is no burden of reclassification involved.

The Science Reference Library, by contrast, was a postwar creation with its nucleus in the nineteenth-century Patent Office Library, and although its book acquisition programs have been extended to embrace all copyright deposit items in all of its disciplinary fields, it still has the particular task of supplying the requirements of industrial research and the patents community—a task which it brought with it from its origins as the Patent Office Library and the National Reference Library for Science and Invention. Furthermore, it has put a large investment of professional energy over the last ten or more years into developing a special classification of its own from an earlier form used in the Patent Office. For convenience we may call this system the SRL scheme. Its collections are arranged by this scheme, and so are its subject catalogs.

The problem here, then, is that the Science Reference Library sees itself as continuing to act as a discrete, “mission-oriented” library with its own role, its own identifiable clientele, and its own classification scheme as one of the tools by which it serves that clientele; the fulfillment of this mission is seen by its staff to require little interaction with its future next-door neighbor, the British Museum Library. What is to be gained by reclassifying SRL for the sake of uniformity with the BML next door?

There are two answers to that question. One is the proposition that the use of a single classification is cheaper overall than the use of two or more—especially when the single classification is one generated from the third corner of the triangle, the Bibliographical Services Division, and is salable, so to speak, to other libraries in the country and overseas. One drawback of the SRL’s own scheme is that it is not used by any other library, and that its addition to BL’s data base would be merely an additional expense, yielding benefit to none but SRL itself.

The other answer to the question is that a single classification for both libraries of the Reference Division will provide flexibility and economy in demarcating the spheres of responsibility for each of the two libraries, as new disciplines and cross-disciplinary literatures emerge in the future. It will allow a uniformity of approach to the changing needs and interests of British Library users in general.

Neither of these answers takes account of the quality of particular classifications, however. If one classification scheme is as good as another, the argument for standardization is simply one of cost and administrative
economy, and the value of classification as a professional method is in some ways diminished. If you regard classification as one of the most important elements in good library service to a particular clientele, you must give weight to the argument that what is gained financially and administratively in a change of classification may be lost in service to readers, if the new classification is less effective than the old one in meeting the needs of those users.

It was not for the Working Party to determine whether SRL was to continue its "mission-oriented" role indefinitely, or alternatively to change its identity into a kind of scientific twin of the BML. What we could do, and did, was to examine the case that its present classification was significantly more effective for the control of scientific literature as a whole than any of the general classification schemes, which, being suitable for the BML, might also be used by SRL as an alternative to its own scheme for shelf arrangement. A study was carried out for us by Aslib of four general classification schemes: DDC, UDC, LCC and the Bibliographic Classification (BC). The last of these four was, of course, familiarly known in its original form by the name of its inventor, Bliss. What we studied was the preliminary schedules of the new revised version developed by Mills and others in England as a faceted classification.

The four schemes were compared with each other and ranked in terms of seventeen criteria which in summary may be grouped as follows: (1) effectiveness as classifications, as evidenced, for instance, by provision of helpful collocation, level of specificity, up-to-dateness, notational qualities and searchability; (2) ease of use by a classifier; (3) availability of schedules; (4) frequency and extent of revision; (5) mechanisms established for maintenance and revision; and (6) extent of present use by libraries and bibliographic services. Judging by most of the criteria related to effectiveness as classifications, and with particular attention to the most recent thought on classification principles (as exemplified, among others, by the Classification Research Group), none of the four schemes was rated very high. We felt that LCC was the scheme that probably met fewest requirements, and BC possibly the most. However, BC has yet to be published in its revised form, and we were assessing it largely on the basis of its authors' claims; the judgment is thus very tentative. DDC and LCC were both rated high for criteria related to ease of use, availability of schedules, and extent of use by libraries as a whole. In addition, DDC was the only scheme of the four that scored high for the criteria related to frequency of revision, and to mechanisms for maintenance and revision.

The investigation closely examined all of these questions, and took account of previous studies, such as the ALA Resources and Technical Services Division Classification Committee's "Statement on Types of Classifi-
cation Available to New Academic Libraries.”¹ The detailed analysis is to be found in a technical memorandum prepared by Aslib for the Working Party, and this will be included among the appendices of the published report. At the risk of considerable oversimplification, I can attempt here only to summarize our findings on the four schemes as candidates for the role of arranging one-quarter million books on the open shelves of the British Museum Library:

BC—This was potentially the most progressive and satisfactory system for future needs. It is being developed in its revised form by British classificationists, and is thus distinctive and symbolic as a significant national contribution to modern classification, in advance of all others. However, against this must be balanced the unknown and untried performance characteristics of the new BC and the lack of any determinate policies with respect to all the other criteria related to availability of schedules, revision machinery, and use by other libraries.

LCC—This is a series of classification schemes used by a significant number of academic and other research libraries, whose general objectives and collections have more in common with the British Library than do those of the British Library with those of other libraries. LCC numbers are also available from the MARC record. Nevertheless, LCC rates low on most modern criteria related to effectiveness as a classification scheme and/or revision mechanism; it was particularly felt that the scheme offered the least possibility for a necessary British Library participation in long-term development and revision in accordance with British needs.

UDC—In many ways, UDC is the most important scheme in the interests of international standardization, particularly as it is widely used in Europe. On the other hand, it is generally considered seriously out of date and in need of drastic revision, and its future is uncertain. If the British Library adopts UDC, it will be necessary for it to become closely and positively involved in schedule development, and probably to make some financial investment in it. In addition, from the British Museum Library user’s point of view, UDC has some disadvantages in its notation, which is designed primarily for the arrangement of document citations, and not for shelf arrangement.

DDC—This is the scheme that emerged as the most likely candidate for the British Museum Library, not so much from its positive merits as a classification (although it was thought to have no fewer positive merits than any of its rivals) as from the relative absence of drawbacks. Its practical advantages were seen to be: (1) it is more widely used in Great Britain than any other scheme; (2) its schedules and index are complete, widely available, and reasonably easy to use; (3) it is already being applied to British copyright
materials, as well as to a substantial proportion of other BL acquisitions; (4) there are well-established procedures for its maintenance and revision; and (5) indications were given to us by British representation on the Editorial Policy Committee, and by recent policy statements from that body, that British and European requirements can be effectively input into the revision machinery in the future.

Classification of the SRL

From these conclusions relating to BML, we were left logically with only three possibilities to investigate with respect to the Science Reference Library. These were: (1) to retain the present SRL classification, (2) to replace the SRL scheme with UDC, or (3) to replace the SRL scheme with DDC. The debate on the relative merits of DDC, UDC and the SRL schemes tended to revolve around three factors: collocation, specificity, and class occupancy. A study carried out by Aslib compared the extent to which UDC/DDC and the SRL schemes collocated works on related topics, and judged that the schemes were roughly equal in this respect. On the other hand, another study by the staff of the Science Reference Library concluded that the SRL scheme provided significantly better collocation than DDC for searches in the field of technology.

Then Aslib undertook a comparison of the specificity of the SRL scheme, the medium edition of UDC, and DDC in three subject areas. Only 54-59 percent of SRL classes had corresponding classes in UDC, while the figures for DDC were 36-38 percent. However, it was felt that it would be wrong to conclude from these results that the medium edition of UDC was less specific than the SRL scheme. A more detailed analysis of the situation, based on the class physical chemistry, showed that whereas only 58 percent of the SRL classes could be located in UDC, only 32 percent of the UDC classes had counterparts in the SRL scheme. Thus, since the overlap between the classifications was much less than might have been expected, there were no clear grounds for concluding that one of the schemes was more specific than another.

At this point, the SRL staff introduced the concept of class occupancy, to be measured as the number of documents filed at a single classmark; they defined an overcrowded classmark as one at which more than twenty documents were filed. Two studies were made of class occupancy and overcrowding. These can be compared only in very general terms, because of the different document samples used, and at this level of comparison they appeared to produce conflicting results. The first study by SRL found that, for three selected subject areas, between 3 percent and 24 percent of the SRL
collection (classified by the SRL scheme) fell within an overcrowded class. If DDC were to be used, it seemed that this percentage might rise to 56-85 percent. However, this study had unfortunately omitted the facilities for synthesis in DDC, and the second study carried out by the Systems Development Branch of BL analyzed the effects of lifting this limitation, thereby allowing DDC a greater degree of specificity. This second study analyzed a complete set (as much as could be obtained) of all statistical mathematics entries classified by the eighteenth edition of DDC in BNB and LC MARC tapes. The set of 911 records produced 225 unique classmarks. Only 2 percent of the DDC classmarks were found to be overcrowded, but 36 percent of the sample fell within an overcrowded class. A further study of the effect on the full sample of eliminating those items which would not meet the criteria of a postgraduate selection policy would be necessary to measure the realistic level of class occupancy which would result, but it is a safe assumption that such an elimination would reduce significantly the numbers of documents in overcrowded classes. Therefore, from an SRL viewpoint, and again at the risk of oversimplification, the arguments for and against the three schemes studied can be stated as follows:

UDC—The defects of UDC are the same for the Science Reference Library as they would be for the British Museum Library. Nevertheless, as an admittedly science-oriented general classification, UDC is by tradition the first choice for scientific and technical libraries of many kinds; some beneficial spinoff in the direction of these other libraries might be expected from its adoption by SRL and its consequent inclusion in the central bibliographic record. If it were adopted by SRL, there would be a strong case for it also to be adopted by BML. As with any other classification not already in use at SRL, there would be on the debit side the cost of reclassifying some or all of the present SRL stock.

DDC—Again, those merits of DDC indicated for BML requirements would apply also for the Science Reference Library. The adoption of DDC would have the additional advantage of reducing the current work load, since a proportion of SRL's intake would be received with DDC classmarks already assigned. DDC shares with UDC the advantage of being an acceptable classification in principle to form the basis of a unified approach to shelf arrangement within the two libraries of the Reference Division as a whole.

On the debit side, as compared with the SRL scheme, there are the costs of reclassifying to DDC some or all of the SRL stock and, in contrast with UDC and the SRL scheme, the relative absence of recognition of DDC by scientists and technologists as a classification particularly well suited to their needs.
SRL Scheme—The SRL scheme has been tailored to the library’s requirements during its primary period of growth as the National Reference Library for Science and Invention. Being an “in-house” scheme, it is entirely under the control of SRL staff, and may be modified at need to reflect changing user requirements, changing acquisition policies or changes in the literature as they occur. The retention of the SRL scheme would avoid the immediate cost of reclassifying some or all of the present stock. In the long run, however, some effort would be required to keep it up to date, that is, to avoid the situation of accelerating obsolescence that befalls all “homemade” classification schemes when their originators depart, or that has arisen to a lesser extent with UDC. This effort would not be offset by cost savings in other ways, although it may be that these savings would be relatively small.Retention of the SRL scheme would also involve the addition of extra SRL classmarks to certain categories of material within the BL data base.

The picture which emerges from studies of the suitability of DDC, UDC, and the SRL scheme for the Science Reference Library collections was thus unclear. The evidence we gathered did not demonstrate a clear superiority of any one classification scheme over another in terms of collocation, specificity or class occupancy; any decisions for changing from the present SRL scheme will have to be made on other grounds.

Two main conclusions followed from the whole investigation of classification schemes. First, if the British Library’s two references libraries are to be regarded as a pan-disciplinary collection with a single classification, the choice for shelf arrangement appears to lie between UDC and DDC. Of these, UDC has a wider international authority (in Europe at least), and a more widely participative process of schedule development; but, as far as the national library community is concerned, DDC predominates.

Secondly, if the British Library Reference Division is to be regarded as two separate collections with a fairly permanent demarcation between them, different classifications for the two collections can be considered. As there is no intention of carrying over the existing pressmark system into the new BML Reading Room, the best choice of existing schemes there would appear to be DDC. For SRL there is no obvious best choice. The advantages of retaining the in-house scheme must be weighed against the long-term, overall advantages of changing to DDC.

We made two further points about the Science Reference Library. The first was that considerable effort had gone into the creation and implementation of its special in-house classification scheme over the previous ten years. However persuasive the arguments of cost-effectiveness might be, the
abandonment of the SRL scheme might well be seen by the specialist staff of the Science Reference Library as an unfavorable verdict on their contribution to the library’s work. It is difficult to weigh this factor, but there is no doubt that it must in some way be inserted in the equation of the decision-making process.

The second point we made perhaps counterbalances the first: the Science Reference Library today puts the greatest emphasis on meeting the needs of those who visit it, and thus rates classification for shelf arrangement high among its professional methods. Looking into the future, we may expect that the SRL will develop rather more into a central component of the developing national and international networks of scientific information transfer, and that it will be better able to perform this function if its classification and indexing systems match those of other components in the network, and follow the lines of the development being pursued within UNISIST. This requirement is likely to outweigh shelf arrangement as an institutional priority, and to reduce the validity of such measures as “class occupancy” in the determination of classification policies.

We had little help from the published literature concerning the determination of unit costs. Most of the unit costs quoted were so dependent on the particular circumstances and environment in which they arose that no useful generalizations could be made. There was a considerable amount of internal information in various forms, and the report’s appendices included analyses and inferences drawn from them. However, they must be regarded as very tentative, since they were derived from data concerning stock sizes and rates of growth that may already have been out of date at the time we examined them.

We had hoped to get more accurate measurements by means of planned diary surveys of operations in various parts of the British Library, but unfortunately these had to be suspended. From what was available to us, we were able to reach certain preliminary conclusions:

1. PRECIS, as well as being more effective, would be significantly cheaper than the present manual system for compiling the British Museum Subject Index.
2. A single classification system for the whole of the British Library would be cheaper to operate than two or more systems.
3. LC Classification probably costs less to apply than any other of the general schemes.
4. DDC appears to cost less to apply than the SRL scheme.

Beyond these, there were few positive statements to be made.
Postscript

Since this paper was presented, the British Library has published the Working Party’s report,\(^2\) with a preface by the Director-General of the Reference Division stating that BL “accepts the recommendations... in general.” Specifically, DDC is accepted “insofar as a single scheme proves to be necessary... [and] will be adopted immediately for certain open access collections.” At the Science Reference Library, however, adoption is to be postponed “until the future accommodation pattern becomes clearer” and in the interim “all current intake will... be classified also by DDC” so as to “minimise the work involved in transferring at a later stage to a DDC arrangement, if this proved to be the best decision for a unified collection.” The costs of delaying the SRL decision, says the Director-General, “though not negligible, are capable of being accommodated.”

REFERENCES