

# The Role of WorldCat in Resource Sharing

By

Anna H. Perrault

## Abstract

The 30<sup>th</sup> anniversary of WorldCat was celebrated in 2001. At that time, there were 45 million records with over 750 million location listings, spanning over 4,000 years of recorded knowledge in 377 languages. In the anniversary year, a bibliometric study was begun under the auspices of an OCLC/ALISE research grant. A 10% systematic random sample of the database was analyzed utilizing the OCLC iCAS software to profile the monographic contents of WorldCat by type of library, subject and language parameters. The profile reveals the extent of global publication made accessible through the OCLC international network. Several findings of the study can be examined as possible barriers to successful cooperation in collection development and resource sharing. One of the major problems analyzed in the study is the timeliness in the availability of bibliographic records for current publications. This paper explores the feasibility of using WorldCat as a cooperative collection development tool as well as additional measures which might be derived from analyzing bibliographic records. The results can be used to stimulate discussion on the role of WorldCat as an international resource on the universe of publication available for research and resource sharing worldwide.

## Introduction

The history of cooperative collection development is familiar to those engaged in the acquisition of information materials and the building of library collections. Most histories of CCD in the United States use the Farmington Plan (1942) as the beginning point of the CCD movement in the 20<sup>th</sup> century. The introduction of electronic resources at the end of the 20<sup>th</sup> century brought new formats and new challenges to cooperative collection development. There was also a renewed intensity of effort as the effects of price inflation in published works in the 1970s and 1980s made CCD an ever more desirable, if elusive goal. There were several conferences and volumes devoted to cooperative collection development at the end of the 1990s. In just a few years, there was an issue of *Library Trends*,[1] several monographs, (*Collection Management for the 21<sup>st</sup> Century*,[2] *Cooperative Collection Development: Significant Trends and Issues*,[3] *Creating New Strategies for Cooperative Collection Development*,[4] and conferences, the CRL 50<sup>th</sup> Anniversary conference at Aberdeen Woods in November 1999, among them, all devoted to resource sharing and cooperative collection development. At the first Aberdeen Woods conference, one of the goals was to develop a new agenda for CCD. There were a number of papers at that conference which focused on new models of CCD in the electronic environment. There was also an international focus in that several papers concentrated

on the necessity for global cooperation in the 21<sup>st</sup> century, that the Internet has fostered the concept of a “global information commons.” This paper is based upon that idea of global cooperation in collection building and resource sharing and offers suggestions for overcoming barriers to CCD. It utilizes findings from the author’s “Global Collective Resources: A Study of the Monographic Bibliographic Records in WorldCat,”[5] which is published on the OCLC Office of Research website. ([www.oclc.org/research/grants/reports/perrault/intro.pdf](http://www.oclc.org/research/grants/reports/perrault/intro.pdf))

The first part of the paper reports findings from the WorldCat study and poses suggestions arising from those findings. After that the standardization of measures for CCD and the role of databases in facilitating implementation of those measures are discussed.

## Global Collective Resources

The WorldCat study was conducted by extracting a 10% systematic random sample of the monographic bibliographic records in the WorldCat database in February 2001. The sample contained 3,378,272 usable records. The first stage analysis was conducted via the OCLC iCAS product which was developed by WLN to provide collection analyses and peer comparisons of library collections. The WorldCat study profiles the monographic bibliographic records in the database by parameters of imprint year, subject, language, adult/juvenile, and unique titles/overlap. Libraries were divided into five type of library groupings: research, academic, public, special/other, and school. The sample was of the WorldCat database and thus includes the international membership of OCLC.

Among the findings of the study, are that the universe of materials under bibliographic control in WorldCat show a high level of diversity of resources with 53% of records in the analysis having only one library location symbol. The unique records in the WorldCat sample are almost evenly divided between records with call numbers and records without usable call numbers. The research libraries have 63.5% of total records as unique records. The other four type of library groupings (academic, public, special, and school) have low percentages of unique titles and higher overlap among the four. The profile of the research libraries grouping by imprint year and subject divisions closely parallels the profile of WorldCat.

Of all the records in the study, approximately 65% are for English language monographs with 35% being foreign language publications. The six foreign language groupings analyzed in the study—Chinese, French, German, Japanese, Russian, and Spanish—together account for 67.5% of all foreign language records in the sample.

One of the most significant findings was identified at the outset during the preparations for the iCAS analysis. It was apparent that over one-third of the records in the sample did not have usable call numbers for subject analysis. Since the call number, Dewey, LC or NLM, is used to perform subject analysis of the records, over one third of the records could not be used in the subject analysis. Analysis of the records without usable call numbers to determine the characteristics of that subset of the sample became an added dimension to the study.

A major finding of the WorldCat study germane to the topics of resource sharing and CCD is that for the research, academic and public library groupings, by nearly every measure, the number of bibliographic records declines rather precipitously in the latter or most current imprint years in the study. The number of records with call numbers declined 72 percentage points between 1992 and 2000, while the number of records without usable call numbers declined 88 percentage points in the same time frame. The decline in the number of records

annually in the 1990s is most acute in the research libraries aggregated collections. The major differences between the research libraries and the academic libraries groupings are in the records without call numbers. Most of the measures show a peak in the early 1990s, around 1992 and 1993, and the number of records declines from that point forward. It seems reasonable to assume that acquisitions and cataloging lag combined could be the reason for the decline in the number of current records. To look at the decline and also to further examine the records without usable call numbers a special analysis was run of those records in the 10% sample which had ISBN numbers. A look at the findings from the ISBN analysis illustrates the downtrend in the number of records in the 1990s.

Figure 1, “WorldCat Records with ISBN Numbers, 1990-2000,” shows the trendlines for those records with ISBN numbers published within the last decade in the study. The ISBN records are analyzed by categories of English language titles, both with and without call numbers, and all non-English titles, with and without call numbers.

It is easy to see that the category with the lowest number of ISBN records, the bottom line in Figure 1, is that of English language records without usable call numbers. Titles with ISBN numbers are mostly from mainstream publishers and are more likely to be cataloged and classified. Hence, it stands to reason that the category of English language records without usable call numbers would have the lowest number of records with ISBN numbers. In fact, “mainstream” publications can be operationally defined as those titles with ISBN numbers.

The next line up from the bottom of the graph is that of non-English (foreign language) titles without call numbers. Not all foreign language titles would have a Dewey, LC or NLM call number, so it is understandable that more foreign language than English language titles with ISBN numbers did not have usable call numbers in the records. The third line from the bottom in Figure 1, is a total of the lowest two lines. Thus, those records with ISBN numbers but no usable call number formed the smallest grouping of records with ISBN numbers in the sample.

The top three trendlines in Figure 1 track those bibliographic records with both ISBN numbers and call numbers. The lowest of the three is foreign language records with call numbers. The number of English language titles with both call numbers and ISBN numbers is higher than the number of non-English titles. Again, not a surprising finding. It can be seen that these two lines for English and non-English records with both call numbers and ISBN numbers do not have exactly the same pattern for the imprint years shown. The English language records with call numbers begin to trend downward in 1992, but the decline is slight and the number of records in 1990 and 1998 are very close to the same number. On the other hand, the foreign language records trend upward and remain in the same range until after 1997, but then drop precipitously. From these data, it appears that English language records with call numbers and ISBN numbers are added to the database in a more consistent and timely manner than the other categories of records analyzed. The decline for records with ISBN numbers is most severe in foreign language records with call numbers, and in all records without call numbers. Thus it appears that mainstream publications in English are the only category of materials added to the database in a timely manner. So we can see that the records which have neither ISBN number nor call number, those materials which are more unique and really establish the diversity of the universe of resources, are added in an incremental manner over a longer period of time.

There are some indications from the WorldCat data analysis that the national libraries and large research libraries internationally which began contributing records in the 1990s are responsible, in large part, for the foreign language records in WorldCat in the last decade in the

study. The Association of Research Libraries initiated the ARU/ARL Global Resources Program in the 1990s to test the feasibility of increased reliance upon non-U.S. libraries for access to foreign material not found in North America. At the first Aberdeen Woods conference, several papers urged more cooperation internationally to depend upon research libraries abroad to acquire the publications from their respective countries. [6] The WorldCat study results are encouraging that libraries abroad do appear to be contributing the bibliographic records for publications from their own country to the WorldCat database.

The availability of bibliographic records for recent publications, especially for foreign language publications, might possibly look better if vendor records had been included in the WorldCat analysis. In the 1990s, many of the major book vendors in the U.S. and Europe began transmitting their database records to OCLC. These records are not included in the 10% sample because they would not have library location holdings even though they may have call numbers. Because the WorldCat study analyzed the sample records by type of holding library, one of the criteria for inclusion in the study was that the records had to have at least one library location symbol. For resource sharing it is necessary to have library holding symbols attached to the bibliographic record. Thus, the findings from the WorldCat study show the availability of titles for resource sharing.

In order to discuss the “Role of WorldCat in Resource Sharing,” the implications of these findings for resource sharing and coordinated cooperative collection development are considered next.

## **WorldCat and Resource Sharing**

In reviewing the present role of WorldCat and implications for the future, it is useful to give some historical background on OCLC and ILL. BD, that is, Before Databases, the process of interlibrary loan was one of detective work and intuition. It was necessary to verify the existence of an item before one attempted to secure the item from another library. The tools for verifying titles were printed catalogs and bibliographies. These tools were mainly of two types: they were by country from the book trade or they were produced from the cataloging records of libraries. In the United States we had *PTLA*, *Books in Print*, *Cumulative Book Index*, from the book trade and the *Library of Congress Union Catalog*. There were the *British National Bibliography* or *British Library Catalogue*; *Bibliographie de la France* or the *Bibliothèque Nationale Catalogue*. The national bibliographies of the most industrialized countries often swept theses and dissertations, government documents, reports, and other grey literature into the comprehensive of “everything published in” a particular country. Before databases, it was much easier to verify the existence of a title than to actually obtain a copy because not all of the tools had library location symbols. Nonetheless, bibliographic coverage was quite comprehensive, at least for the industrialized countries. But only large research libraries owned the full complement of the bibliographies and catalogs needed for the verification work. This is one example of how BD a scholar could be disadvantaged if they were not working in a research institution.

When the OCLC Catalog began in 1971 with a few thousand records, the task was to build up a cataloging database by adding current publications but also records for the existing universe of retrospective publications. After 30 years, the database can now be said to have reached critical mass with respect to the universe of publication. But unlike the printed

bibliographies, an item does not appear in WorldCat unless at least one library owns it. With the exception of vendor records, as long as there is a bibliographic record in WorldCat for an item, there is at least one library holding symbol. So in a way, we have an opposite situation than before databases when we had fairly good bibliographic control of what items were in existence whether or not any library held a copy. Now, to most library staff, if a bibliographic record for an item is not found in WorldCat or other databases, that item, for all library practical purposes, does not exist. We have shifted the responsibility of bibliographic control from comprehensive country specific printed bibliographies and other tools not originated by libraries to dependency upon libraries to contribute cataloging records. Which brings us to the implications of the decline in the number of bibliographic records within the last, most current decade of publication.

For purposes of resource sharing, the retrospective wealth, numbers and variety of resources represented in WorldCat provide the where-with-all for over 8 million ILL transactions annually. For resource sharing purposes, WorldCat is where the action is and this action is increasingly international. OCLC has revolutionized the process of resource sharing, but the same cannot be said for CCD. Whereas, the great majority of ILL is either for retrospective monographs or journal article photocopies, CCD centers around current acquisitions. And we have seen that the universe of publication is not fully represented in records for recently published materials. One of the major purposes of CCD is to ACQUIRE things. CCD is focused on the present and what will be acquired for the future. Given the findings from the WorldCat study and other similar bibliometric studies, it seems there are problems for which we need to seek solutions.

The next section contains suggestions for approaches to the problems pointed up by the findings of the WorldCat study.

## **Cooperative Collection Development**

The problem of contributing current bibliographic records to WorldCat was addressed in 1999 when the OCLC Users Council adopted a resolution on “Shared Commitments to the Principles of Cooperation.” The thrust of the resolution was to re-affirm and strengthen OCLC’s commitment to cooperation and resource sharing. A letter was sent to all member libraries and networks urging compliance with the responsibility to contribute all current, Roman-alphabet cataloging records and holdings to OCLC. The letter stated that “these actions, fully supported by Users Council, are intended to safeguard and strengthen WorldCAT as the pre-eminent, international union catalog and the foundation for global library collaboration.”[6] While it has only been three years since the User’s council resolution, as of the time of data extraction in February 2001 for the WorldCat study, it does not appear that there has been an effort on the part of member libraries to increase the availability of bibliographic records for recently published materials.

Indeed, there is evidence to conclude that cataloging has become a diminished activity in research libraries. Wilder’s study of “New Hires in Research Libraries,” found a 45% decline in the hiring of new catalogers between 1985 and 2000.[7] This finding certainly dovetails with the declines in the cataloging of new publications, especially those without ISBN numbers, found in the WorldCat study.

The findings of the WorldCat study indicate that current imprints with ISBN numbers are not where the problem lies in the addition of records for recently published titles. The drop in the

number of records for recently published titles may be as much a problem of becoming aware of the existence of those titles as of timely cataloging. This is the old chicken and the egg syndrome. How do we find out about the existence of publications? We check WorldCat, vendor databases and other electronic indexes and databases. But how do records for publications get into the database? This points up how dependent we are upon the electronic bibliographic databases. While there was a drastic decline in the hiring of catalogers, there was an encouraging 61% increase in the hiring of subject specialists, leading one to hope that there are professionals to hunt for appropriate titles in addition to the obvious ones with ISBN numbers.[8])

### Suggestion Number one

*CCD agreements should have the responsibility for timely cataloging and classifying of materials as well as the acquisition of them.*

It is pretty useless to resource sharing partners for a library to claim to have the definitive collection of materials in a certain subject area if the bibliographic records to identify the items in those collections do not exist. Acquisitions is only the first step in CCD. Bibliographic control is the other half of the equation. Perhaps not enough attention has been paid to the supplying of bibliographic records as a corollary of cooperative collection development agreements.

### Suggestion number two

*Creation of incentives for cataloging current materials.*

When OCLC was first founded, a great deal of attention was paid to building up the database. Policies were set to give incentives for adding bibliographic records for retrospective resources. Now perhaps incentives are needed to both acquire current materials **without ISBN** numbers and catalog and classify those items. Maybe adding a record for a current item without an ISBN number could be free? Could it actually become profitable to add records for certain kinds of materials? But only if added, say within two years of publication date. Would it be possible to create a competitive mentality whereby instead of waiting for a record to show up in the database, librarians actually would vie to be the first to enter a record?

In 2001 at the ALA Annual conference in San Francisco Phyllis Spies, Vice President, Worldwide Library Services for OCLC gave a presentation entitled, “Key Barriers to International Resource Sharing and OCLC Actions to Help Remove Them.” One of those barriers was described as the “lack of a critical mass of online metadata for the world’s libraries,” and one of the solutions to overcoming that barrier was the suggestion to provide “contribution” credits to encourage shared cataloging on OCLC. Other suggestions were to load more non-U.S. library catalogs into WorldCat and to link WorldCat to more country/region based union catalogs.

The point here is that in spite of all these efforts, it appears that a large proportion of the records in WorldCat within the most current five imprint years are for mainstream publications with ISBN numbers. Perhaps one of the problems is that only these materials are being acquired on a current basis. It may be that more attention needs to be focused on the timely identification

and acquisition of local materials and other publications which enrich the diversity of bibliographic databases.

### Suggestion Number three

*Can database scans be run for records in WorldCat without library holdings to identify titles still in need of an owner?*

The bibliographic records contributed by vendors and other pre-publication cataloging might be useful in CCD for identifying titles no library owns. In other words to get maximum benefit from the vendor contributed records it would be useful to know, say, three years after publication date, those records still without at least one library holding symbol, that is, those records which have not been converted into real bibliographic records because no library owns the title. This would be a measurement of what is not being purchased. Creating lists of this type might be a real challenge for CCD arrangements because those titles not owned would become... who's responsibility?

These suggestions are made to encourage the acquisition, cataloging, and classifying of recently published materials, with an emphasis on titles which are not mainstream publications that will add to the variety of resources represented in WorldCat. There are other ways in which the records in WorldCat can be utilized in resource sharing and CCD. One of these avenues is in the provision of data for measuring the impact or success of consortial agreements.

### **CRL Working Groups**

Several working groups were formed out of the first CRL Aberdeen Woods conference in 1999. One of these is the CRL/Greater Western Library Alliance Working Group for Quantitative Evaluation of Cooperative Collection Development Projects. The measures drafted by this group include a number of performance measures for which the data need to be produced from library systems management reports, consortial data reports, or databases such as WorldCat. The point is create standardized measures and to monitor changes in those measures over time. Success can only be judged in improvements in the number and variety of resources being provided to users. Routine management reports can be designed to provide data for those measures which track the diversity of resources, currency of resources, retrospective depth of resources, etc.

A few of these measures suggested by the CRL/GWLA Working Group for Quantitative Evaluation of Cooperative Collection Development Projects are:

- The number of non-serial and serial titles by subject and time period.
- The number of unique titles by subject and time period.
- The mean number of holding libraries per title (on the average, how many libraries in the consortia own each title)
- The median age of collections by subject.

The OCLC iCAS product either already provides the means to assess consortial progress or could be adapted to provide those measures from WorldCat data. ICAS (Interactive Collection Analysis System) is a component of the OCLC ACAS Collection Analysis Service. The iCAS product provides customized age and content analysis of library collections with title overlap and gap analyses.[9] There have been several state assessment projects conducted utilizing the iCAS product. It is also being used abroad. For example, CURL (the Consortium of University Research Libraries in the British Isles) began a collection analysis project using the iCAS software in June 2001. Besides using the product within a consortia, the possibility also exists of comparing the aggregated holdings of different consortia across standardized measures.

There are now a number of library systems which have incorporated such measures as median age or number of unique titles into the system management reports. Consortial databases have also been programmed to provide the number of libraries owning each title and identifying unique titles. If we had a sufficient number of consortia reporting standardized measures, these measures could become benchmarks for libraries and consortia to utilize in collection comparisons.

Up to this point, the annual report statistics from OCLC have not focused on providing data to assist in CCD, but this could be done. ICOLC could adopt measures and work with the bibliographic utilities to provide these annually as benchmarks. IFLA has for many years had a goal of universal availability of publications. It now seems possible to attain this goal with international databases and international agreements.

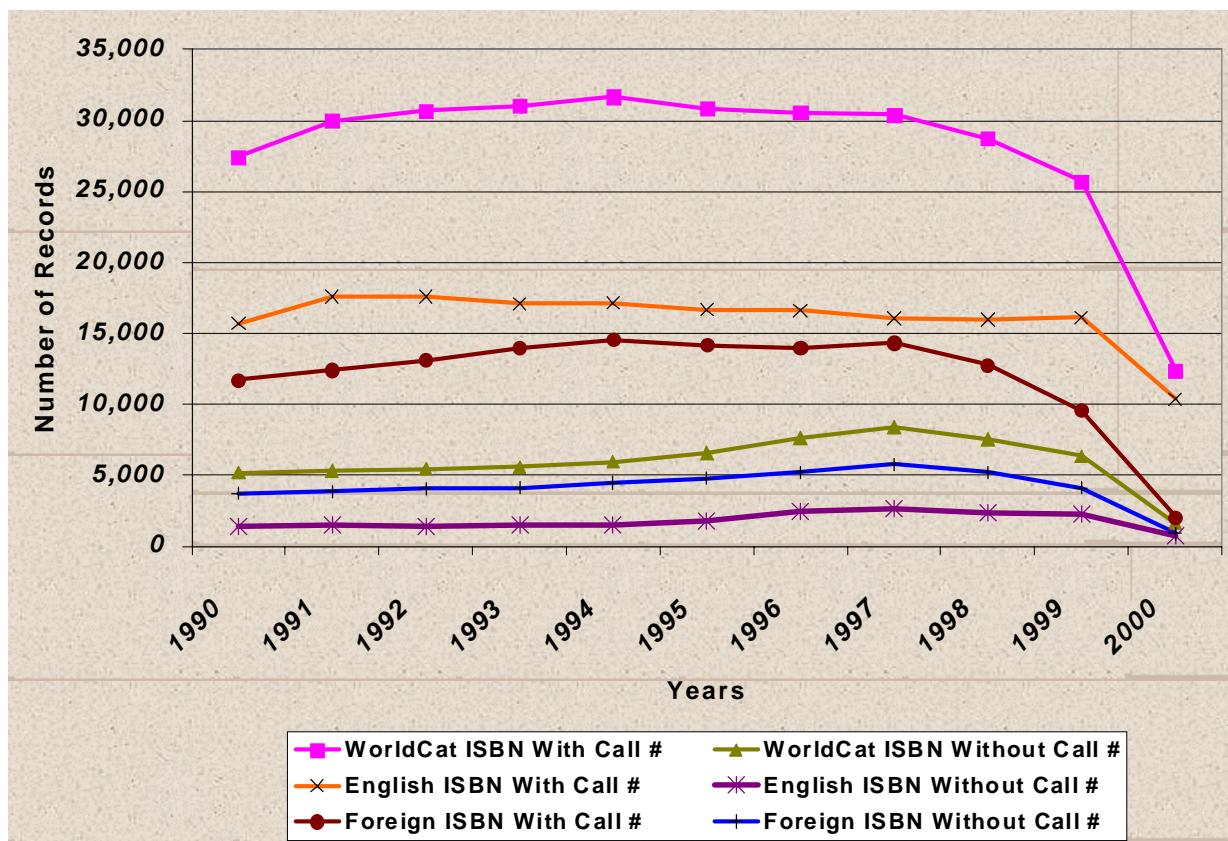
## Conclusion

In an article on the international growth of OCLC in 1998, Phyllis Spies, stated, "OCLC offers the opportunity to facilitate the identification and location of materials on a global basis. OCLC and its member libraries are building the global information infrastructure. This collaborative effort allows people around the world to gain access to the rich scholarly resources of libraries that heretofore may have been inaccessible due to the lack of automation and/or isolation of a country. In the 21<sup>st</sup> century, OCLC and its member libraries have the opportunity to achieve the long established library goals of universal bibliographic control and universal access to publications."<sup>[10]</sup> WorldCAT as the pre-eminent international bibliographic database has a role to play in providing the means for formulating cooperative agreements in this global effort.

OCLC has continued to work with libraries to load national collections into WorldCat. The list of these libraries is becoming very long and the international representation in WorldCat continues to grow. The number of international book vendors and publishers contributing records to WorldCat also continues to grow.

With respect to enrichment of the database with unique resources, it may be necessary for there to be a collective change in acquisitions and cataloging procedures. It is the responsibility of the OCLC membership and those who wish to engage in CCD to ensure that the international universe of publication is represented in WorldCat.

**FIGURE #1**



"WorldCat Records with ISBN Numbers, 1990-2000"

## References

1. *Library Trends* 45 (Winter 1997).
2. *Collection Management for the 21<sup>st</sup> Century*, ed. G.E. Gorman and Ruth H. Miller. Westport, CT: Greenwood Press, 1997.
3. *Cooperative Collection Development: Significant Trends and Issues*, ed. Donald B. Simpson. New York: The Haworth Press, Inc. 1998.
4. *Creating New Strategies for Cooperative Collection Development*, Pts. 1 and 2, ed. Milton T. Wolf and Marjorie E. Bloss. New York: The Haworth Press, Inc., 2000.
5. Perrault, Anna H. "Global Collective Resources: A Study of the Monographic Bibliographic Records in WorldCat," Report of a study conducted under the auspices of an OCLC/ALISE 2001 Research Grant.  
[\(http://www.oclc.org/research/grants/reports/perrault/intro.pdf\)](http://www.oclc.org/research/grants/reports/perrault/intro.pdf)

---
6. Perrault, Anna H. "The Printed Book: Still in Need of CCD," *Collection Management* 24, nos. 1/2 (2000): 119-136; Don Hazen, "Dancing with Elephants: International Cooperation in an Interdependent (But Unequal) World," *Collection Management* 24, nos. 3/4 (2000): 185-213; Phyllis Mirsky, R. Bruce Miller, Karl Lo, "From Farmington Plan to the Pacific Rim digital Library Alliance: New Strategies in Developing International Collections," *Collection Management* 24, nos. 3/4 (2000): 241-250.
7. Wilder, Stanley J. "New Hires in Research Libraries: Demographic Trends and Hiring Priorities," *ARL Bimonthly Report* 221 (April 2002): 1.
8. Ibid., p.2.
9. Information on iCAS is at <http://www.oclc.org/western/products/aca/icash.htm>
10. Phyllis B. Spies. "OCLC Worldwide," *OCLC Newsletter* September/October 1998): 23.

Anna H. Perrault is associate Professor in the School of Library and Information Science at the University of South Florida in Tampa. She has conducted collection analysis of academic library collections and published widely on collection analysis and assessment.