UCLA/OCLC Core Record Pilot Project: Preliminary Report

Sherry L. Kelley and Brian E. C. Schottlaender

This report details the record-creation phase of the UCLA/OCLC Core Record Pilot Project. A total of 384 records were created, consisting of 234 core and 150 full (control) records. Approximately half of the core records were coded K level in OCLC, and half I level. NACO authority work was done for all controlled (name, series, subject) access points in both the core and control records. Core record creation was determined to be significantly faster than control record creation: between 8.5% and 17% faster, depending upon whether learning curves are factored in. The core records created included, on average, 1.52 subject headings and 1.01 name headings each; the control records averaged 2.05 subject headings and 1.59 name headings each. The importance of these differences for access is unclear. Of the 384 records created, 30 core and 15 control records—a total of 45—were subsequently used by 91 institutions within two months of their creation. Of the 45 used, only 7 were modified: 6 core and 1 control. Of the ten modifications made to these records, only two involved the addition of controlled access points. With OCLC support, UCLA will continue to gather use and modification data for a year.

The University of California at Los Angeles (UCLA) University Research Library (URL) Cataloging Department, in cooperation with the OCLC Online Computer Library Center, Inc. (OCLC) and with the approval of the Program for Cooperative Cataloging (PCC), carried out a core record pilot project for monographs from December 1994 to April 1995. Core level cataloging is a national standard developed by the PCC as a means to increase the pool of usable original cataloging records in national and international bibliographic databases. The standard defines a set of data elements essential to cooperative use of the catalog record and requires that all access points be supported by authority records in a national authority file (figure 1).

The purpose of the UCLA project was threefold. First, to follow up on and expand the earlier Cornell study (Cornell

SHERRY L. KELLEY is Head, Cataloging Services Department, Smithsonian Institution Libraries (e-mail: skelley@sil.si.edu). BRIAN E. C. SCHOTTLAENDER is Associate University Librarian, Collections & Technical Services, University of California, Los Angeles (e-mail: bschott@library.ucla.edu). The authors acknowledge, with thanks, the participation of the following UCLA catalogers in the Core Project: Caroline Holt, Carolyn Loeffler, Renee McBride, Nancy Norris, and Frank Peterson. The authors also wish to thank the OCLC Online Computer Library Center, Inc.'s Karen Calhoun for her invaluable assistance. Manuscript received September 1, 1995; revised March 1, 1996; accepted for publication March 12, 1996.

FIXED FIELD VALUES:
Code fully.

020,$a (ISBN):
If present on item

040 (Cataloging source)

042 (Authentication code)

050,082,086, etc.:
Assign at least one classification number from an established classification system recognized by USMARC.

1XX (Main entry):
If applicable

240 (Uniform title):
If known or readily inferred from material being cataloged.

245-300 (Title page transcription through physical description):
Describe fully, using all data elements appropriate to the item described.

4XX (Series area):
Transcribe series if present.

5XX (Note fields):
Minimally, include the following if appropriate:
500: Note for source of title if not from t.p.
505: (Contents note) For multi-part items with separate titles
533: (Reproduction note)

6XX (Subject headings):
If appropriate, assign from an established thesaurus or subject heading system recognized by USMARC at least one or two subject headings at the appropriate level of specificity.

7XX (Added entries):
Using judgment and assessing each item on a case by case basis, assign:
1) a complement of added entries that covers at least the primary relationships associated with a work (e.g. joint authors);
2) added entries to bring out title access information judged to be important.

8XX (Established form of series if different from that in 490 field):
If series is traced, use as appropriate.

Figure 1. Core Record for Print Monograph.

1995), which did not include national level authority work. Second, to test the assumptions that: (a) the creation of core level cataloging will take less time than that required for full-level cataloging; and (b) core records will provide sufficient access to materials through their description and through authorized headings for names, subjects, uniform titles, and series. Third, to provide the PCC with some concrete data for consideration in national program implementation.

The researchers designed the project to gather data about the following questions as well:

1. Subsequent to their creation, how often are core records used by other libraries for cataloging purposes?
2. When they are used, how many modifications are made to them?
3. What kinds of modifications are made to them?
4. Are core records coded level I used differently from those coded level K? (OCLC defines level I cataloging as full-level cataloging input by OCLC participants, and level K cataloging as less-than-full cataloging input by OCLC participants.)

OCLC compiled information to address these questions by using tracking software to record use and modification of the records. These data will continue to be gathered for a year.

Project Scope

Material selected for the project had to meet specific criteria. They had to be monographic works in the roman alphabet script, without full-level cataloging copy in OCLC. Further, the following were excluded:

1. Items requiring new series authority records, or changes to existing LC Name Authority File (LCNAF) series authority records;
2. Items targeted for the UCLA remote storage facility, which do not require classification or subject headings and often receive only minimal cataloging; and
3. Individual belles lettres, which are already very easy to catalog because, at UCLA, they do not receive subject headings.

Project Description

A total of 384 records were created over the course of the project, consisting of 234 core and 150 full (control) records. Sixty of these, 30 core and 30 control records, were created between December 1994 and January 1995 as part of a pre-test. OCLC tracking statistics for the pre-test records are included in the tables below, but time statistics are not. The remaining 324 records were created in a six-week period between February and April 1995.

After a one-hour training session, the catalogers began creating a preliminary set of "practice" core records, following guidelines developed from the proposed core record data elements (figure 1) listed in *Cooperative Cataloging Council Task Group 4: Standards, Final Report, October 29, 1993* (Reser 1994, 53–60). These practice records are not included in the project data. Following this, two catalogers began the project by creating core records, and two began by creating control records. Upon completion of the required number, the catalogers swapped assignments.

Five catalogers participated in the project, one in the pre-test phase, and four in the project itself. Each project participant cataloged approximately 80 titles between February 28 and April 4, 1995. Of these, approximately 50 were cataloged following core record guidelines and 30 following the full-level standards as defined in *Bibliographic Formats and Standards* issued by OCLC (OCLC, 1995). Catalogers timed the cataloging process for both categories of material and recorded results on log sheets (see figure 2). The log sheets were submitted weekly to the project manager, along with copies of the OCLC records. The project manager then reported the record ID numbers to the OCLC liaison, Karen Calhoun, to track further use and modification by OCLC member libraries.

Titles in both samples were assigned full classification and all headings were given full authority control: i.e., all name, uniform title, series, and subject heading access points were represented by records in the LCNAF or in the Library of Congress subject authority file. Project catalogers created new NACO records if none existed in the LCNAF, and updated any NACO records needing revision. No new subject headings were proposed. The same authority control procedures were followed for both core and control rec-

DATE:

RECORD TYPE:

RECORD ID # & (CREATE DATE)	ENC LVL	TIME SPENT DESCRIP CAT (HR:MIN)	TIME SPENT SUBJ/CLASS	TIME SPENT AUTH	# OF NEW NACOS	COMMENTS

Figure 2. Sample Data-Gathering Form.

ords. This was done to neutralize the impact of NACO record creation.

Ideally, the sample of material would have allowed for the same titles to be cataloged at both full and core level, thus providing a better basis for comparison. Unfortunately, most cataloging departments must contend with meeting the ongoing needs of their communities of users and cannot afford to catalog the same title twice; the URL Cataloging Department is no exception. Nevertheless, every effort was made to select material on the same subject and in the same language, and to divide those as evenly as possible between both the core and control samples.

Although we endeavored to randomize book selection within subject and language areas and between core and control conditions, we did so in the context of actual workflow through the Cataloging Department. Many choices made in this workflow could have affected the placement of books in the core or control sample. While we cannot, therefore, claim that book selection was truly random, we can say that the project took place under actual cataloging conditions—an important point, since any benefits from the information produced by this study are only valuable in the context of those conditions. It is important to note that no special choices—other than the exclusions listed above—were made in this study.

Nevertheless, the magnitude of the differences between the core and control records is very large for the sample size. We concentrate on these differences, what causes them, and what the consequences of them might be for the libraries engaging in core cataloging.

Research design for any study comparing a new activity to an established one must control for time needed to learn the new activity. Otherwise, the learning curve itself will cloud the comparison between the time devoted to the new activity

and that devoted to the established one. In this study, the researchers tried to neutralize or isolate the learning curve data by having catalogers create practice records before beginning the project, and by rotating their assignments. The two "learning" areas were the new set of cataloging guidelines for the core record, and the introduction of the use of a stopwatch. Interestingly, the catalogers reported that use of the stopwatch was the more problematic of the two new tasks. As an aside, there was a third, unanticipated "learning" area that was probably a constant for all four catalogers. This was the implementation of Format Integration, phase I, on January 28, 1995—roughly one *month* before cataloging for the project began. Some of the learning curve data may include the changes occasioned by Format Integration—in treatment of alternative title information, for example.

Because the UCLA/OCLC Core Record Project began before the implementation of explicit coding values for core records, UCLA, OCLC, and PCC worked out an interim set of data elements to include in the records to mark them as core level and allow for their batch retrieval. These are use of the subfield "e" of the 040 field with the term "core" (040 CLU :e CORE :d CLU) and creation of a 500 field with the phrase "core record." All core records created in this project can be retrieved through a "fin nt core record" search in OCLC. The newly defined values for PCC core records for the fixed field and 042 field were not implemented until after phase II of Format Integration. The need to add these two interim elements may have added some unnecessary time to the creation of core records.

Results

On average, core record cataloging is faster than full level. Tables 1 and 2 contain two different calculations for the time differential. In table 1, the mean time for each set of core records (50 per cataloger) is given, along with the mean time for each set of control records (30 per cataloger). The difference shown is that between the average time spent creating a core record and that creating a control record. Negative numbers indicate that core cataloging took less time than control cataloging. In the comparison shown in table 1, core cataloging takes over two minutes less than full cataloging per record, representing an 8.5% time savings. In all four cases, the last set of records was the fastest, whether core or control. This tells us something about learning time, particularly involving the use of stopwatches.

Table 2 contains the average time spent on approximately the last half of the core records (25 to 29 records) compared to that spent on the full set of control records. Given the small sample, and given the information shown in table 1 suggesting that learning time is a part of the total cataloging time for both levels of cataloging (use of stopwatch for control records, use of stopwatch and new cataloging guidelines for core records), the

TABLE 1

Cataloging Time

Cataloger	*N*	Mean Time per Record (Core)	*N*	Mean Time per Record (Control)	Difference (Core Minus Control)
*A	50	22:58	30	18:40	+4:18
B	50	14:09	30	19:28	-5:19
*C	52	35:50	29	30:13	+5:37
D	52	22:18	31	35:25	-13:07
Total	204	23:59	120	26:00	-2.01

*Began with core-level records.

TABLE 2

CATALOGING TIME—FACTORED

Cataloger	Mean Time, Last Half Core Records	Full Set Core Minus Last Half Core	Last Half Core Minus Control
*A	24:04	-1:06	+5:24
B	13:24	+0:45	-5:04
*C	26:47	+9:03	-1:26
D	21:53	+0:23	-13:22
Total	21:32	+2:27	-4:28

*Began with core level records.

researchers included this data as another means to control for learning time when two new tasks had to be learned in the first set (i.e., for catalogers A and C who began the project with core level cataloging). In three out of the four cases, the time spent per record on the last half of the core records was less than the time spent per record on the entire set of core records. Moreover, this half of core record cataloging took approximately 4.5 minutes less per record than full cataloging, a time savings of over 17%.

NACO authority work was the wild card in this Project. Of the three cataloging areas timed by the catalogers—description, subject analysis and classification, and authority control—time spent on authority control fluctuated most dramatically among catalogers. The number of new NACO authority records created for each set of records is shown in table 3A. In addition, these data show that among all four catalogers only a slightly larger percentage of new NACO records were created for core level cataloging.

Table 3B shows the average time spent on authority work per title. The wide disparity in time spent creating new NACO records can be explained in part by the mix of name and corporate headings. Cataloger B created no corporate names in either set of records.

One of the goals of this project was to evaluate the adequacy of access for core records. The guidelines for creating core records call for the assignment of one to two subject headings only. In addition, the guidelines do not require additional headings for authors or titles in a single work when either numbers more than two. Limiting the number of obligatory headings is one of the efficiencies of core cataloging. Do these reductions limit access, however? Table 4 lists the number of subject headings per title in core and control

TABLE 3A

AUTHORITY WORK

	No. of Name Headings / No. of Records		No. of New NACOs		% Hdgs. Needing NACOs	
Cataloger	Core	Ctrl.	Core	Ctrl.	Core	Ctrl.
A	70/50	50/30	29	18	41	36
B	57/50	39/30	23	13	40	33
C	73/52	53/31	30	15	43	28
D	70/52	53/31	30	15	43	28
Total	270/204	191/120	99	61	37	32

TABLE 3B

AUTHORITY WORK

Cataloger	Avg. Time per Title/No. of NACO Headings	
	Core	Control
A	5:24/29	4:48/18
B	2:34/23	3:16/13
C	11:12/17	10:43/15
D	5:42/30	6:07/15

TABLE 4

SUBJECT ASSIGNMENT

Cataloger	No. of Subject Headings / No. of Records	
	Core	Control
A	73/50	47/30
B	72/50	44/30
C	80/52	71/29
D	87/52	84/31
Total	312/204	246/120

records, and table 5 shows total headings assigned, both subject and name, for both sets of records. Control records received an average of 2 subject headings per title, core records an average of 1.5 subject headings per title. Overall, total headings assigned averaged 2.85 for core records, and 3.64 for control records.

A significant result of the project is the number of institutions that have made subsequent use of project records. An OCLC tracking program shows that 91 institutions used 45 UCLA project records within the first two months of their creation. This is a strong argument for a national cooperative cataloging program. Table 6 contains a list of the records used by record number, encoding level, and number of holdings symbols attached. The table also reflects whether or not records were modified during their subsequent use.

It is interesting to note that of the core records that were modified, two were K level and four were I level. This is the opposite of the researchers' expectations. Since the result given is so small, we cannot consider this statistically significant, however, and will need more use-data before drawing conclusions, if any, about treatment of I and K level records by catalogers.

Ten subsequent-use modifications have thus far been made to 7 of the 45 project records used to date—6 core and 1 control. These modifications have included addition or deletion of fields, and changes to field content. Table 7 is a summary of the modifications. Only two directly affected access: in one case, a variant title field (246) was added; in another, a subject field (6xx). Notes (500) and bibliographical references (504) fields were the most frequently modified. Again, the volume of use-data is too small to draw conclusions at this time, but, at this point, modifications that directly affect access are small in number.

TABLE 5

TOTAL HEADINGS ASSIGNED

Cataloger	Core No. Headings / No. Records	Avg. per Core Title	Control No. Headings / No. Records	Avg. per Control Title
A	143/50	2.86	97/30	3.23
B	129/50	2.58	83/30	2.77
C	153/52	2.94	120/29	4.14
D	157/52	3.02	137/31	4.42
Total	582/204	2.85	437/120	3.64

TABLE 6

USE STATISTICS FROM OCLC

Record Number	Encoding Level	Previous Holdings*	Updated Holdings	Modified (Y/N)
Core				
31746834	I	1	6	n
32011958	I	1	2	n
32056423	K	2	5	y
32058347	I	1	2	y
32065038	I	3	4	y
32069303	I	2	3	n
32069434	K	2	5	y
32084096	I	1	2	y
32084786	K	1	3	n
32104590	I	2	10	n
32098426	I	1	2	y
32105059	I	1	2	n
32152523	I	1	2	n
32154317	I	1	3	n
32164001	K	1	2	n
32164506	K	2	3	n
32078179	I	1	2	n
32078252	K	1	2	n
32112499	I	1	2	n
32197837	K	1	3	n
32198861	K	1	2	n
32184169	K	1	2	n
32229000	K	1	8	n
32222827	K	1	2	n
32234035	K	1	3	n
32254730	K	1	3	n
32266559	K	1	3	n
32267359	K	1	2	n
32297924	I	1	3	n
32291731	I	1	3	n

CONTINUED ON NEXT PAGE

CONCLUSION

On average, core record cataloging was faster than full-level cataloging. In the best case, allowing for project learning curves, the time saved was over 17% per record; in the worst case, disallowing learning curves, it was 8.5% per record. While generally confirming the results of the Cornell study, the UCLA/OCLC

TABLE 6 (CONTINUED)

Record Number	Encoding Level	Previous Holdings*	Updated Holdings	Modified (Y/N)
Control				
32048971	I	2	4	n
32053247	I	1	2	n
32054493	I	2	3	n
32104631	I	3	8	y
32104731	I	1	2	n
32104693	I	1	3	n
31923478	I	8	15	n
31873459	I	3	5	n
32128973	I	3	4	n
32164424	I	2	4	n
32164475	I	1	4	n
32164664	I	1	2	n
32173818	I	1	2	n
31785002	I	2	3	n
32187088	I	2	3	n

*Records with "Previous Holdings" exceeding 1 represent those for which either Enc lvl K or Enc lvl 5 OCLC records were upgraded to core level by Project catalogers.

Total records used/created:	45/384
Core records used/created:	30/234
K level core records used:	15
I level core records used:	15
Control records used/created:	15/150
Total holdings updates:	91
Core holdings updates:	60
Control holdings updates:	31
Total number of records modified/used:	7/45
Core records modified/used:	6/30
K level:	2
I level:	4
Control records upgraded/used:	1/15

project researchers extend the earlier study by demonstrating that significant time savings accrue to core record cataloging even when NACO authority work is factored into the equation.

Core records created during the project include an average of 1.52 subject headings and 1.01 name headings each; Control records include an average of 2.05 subject headings and 1.59 name headings each (26% and 36% more than core records, respectively). It remains to be seen how these differences will affect access, if at all. The data on the kinds of modifica-

TABLE 7

SUMMARY OF MODIFICATIONS MADE DURING SUBSEQUENT USE

Field Tag	No. of Modifications
015	1
041	1
043	1
082	1
246	1
260	0
300	0
500	2
504	2
505	0
6xx	1
7xx	0

tions made to project records subsequent to their creation appear to indicate, at this point, that the impact may be small. Thus far, only two access fields have been added to the forty-five records used by other libraries. The researchers will continue to monitor use-data from OCLC.

WORKS CITED

Cornell's core-level record cataloging in RLIN. 1995. *RLIN focus* Feb.: 4–5.

OCLC Online Computer Library Center, Inc. 1995. *Bibliographic formats and standards*, revision 1. Dublin, Ohio: OCLC.

Reser, D. W. 1994. *Towards a new beginning in cooperative cataloging*. Washington, D.C.: Library of Congress, Cataloging Distribution Service.