Condition Considerations: An Inquiry Into Recording Conditions in Consortial Collections for the Purpose of Selecting (and Deselecting) Shared Print Copies

Michael Garabedian
Whittier College, mgarabed@whittier.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>

Part of the [Library and Information Science Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Michael Garabedian, "Condition Considerations: An Inquiry Into Recording Conditions in Consortial Collections for the Purpose of Selecting (and Deselecting) Shared Print Copies" (2014). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284315571>

Condition Considerations: An Inquiry Into Recording Conditions in Consortial Collections for the Purpose of Selecting (and Deselecting) Shared Print Copies

Mike Garabedian, Whittier College

Abstract

Following preliminary discussions about a shared print network among Statewide California Electronic Library Consortium (SCELC) institutions in which he determined that artifactual condition would not be a criterion for retention, the author developed an online survey instrument for the purposes of verification. The survey was utilized in a condition survey of mutually held book copies at eight SELC institutions. More than 3,400 book copies were examined. Findings indicated that although the majority of books are in “good enough” condition for a shared print network, because just 1/3 of the copies have paratextual elements, it is probable that random deselection of books would result in deaccessioning “duplicates” with artifactual value.

Why I’m Here and What I’m Doing

People who know my library is not actually one of the SELC institutions that expressed even a mild interest in shared print and therefore not one of the pilot cohort involved in the first round of collection and use analysis have asked me what a librarian from Whittier College is doing on our consortium’s Shared Print Working Group, anyway. I think the main reason SELC Shared Print Chair Bob Kieft asked me to participate is because in the informal online and in-person discussions about shared print that led up to the formation of our working group, I raised a point I thought important enough to repeat on several occasions. I first broached this subject in an October 2012 listserv posting wherein Bob sought advice from resource sharing network participants regarding his institution’s (Occidental College’s) large-scale deaccessioning plans:

I wonder if anyone knows offhand if, in discussions about recent deaccessioning procedures and/or resource sharing (especially in regards to shared print book repositories), practitioners have given consideration to the potential artifactual value of the printed books in our general collections? . . . [I]f in going forward, just one or two . . . copies [of a given edition] are going to be part of our shared print repositories, won’t we want to make certain that the copies we retain . . . are the most complete documents we can find?

Since in many ways the impetus for our consortium’s shared print efforts can be traced back to that 2012 message to which I was replying, in one sense paying attention to the condition and completeness of shared print copies has been part of conversations about a potential shared print agreement from the first. And since in response to my reply more than a few member librarians said that in thinking about a consortium-level shared print agreement they favored a policy whereby “best copies” would be identified and retained for sharing, I became our Shared Print Working Group’s representative condition person.

Of course, simultaneous to expressing a desire to retain best copies, colleagues were also quick to identify significant obstacles to considering book conditions for the purposes of retention and deselection. Karen Schneider of Holy Names University zeroed in on one of the main impediments to using condition as a criterion for shared print, writing “I agree with ‘best copy’ [but] am thinking that we have at best very limited tools for this (versus online bookselling where noting condition of copies is routine).” Meanwhile, consultant Lizanne Payne expressed others’ legitimate concern that developing tools to assess and record condition beyond a simple kind of yes/no validation, and then deploying such tools, would likely be too time-consuming, and too costly an undertaking, for the purposes of shared print. (K. Schneider and L. Payne, personal communications, October 2012).

At the same time I was thinking about how we might use condition as a measure by which to identify copies to share, I became eligible for a summer research sabbatical that would allow me to test on a small scale the feasibility of the kind of condition analysis I was proposing. In short, my project was to define the physical attributes condition validation would include, but also to undertake the more difficult tasks of developing the procedures by which condition would be assessed and recorded and then actually to put these procedures into practice by assessing the condition of mutually held copies at several SCEL member libraries.

Why It's Important

Because of the ambiguity of the term but also because it informed my survey instrument and analysis, I want to distinguish what I mean by *condition*. For the purposes of shared print certainly it is imperative to identify damaged books or books with missing pages, for example, or to exclude from consideration book copies whose poor conditions might mean a copy would have to be conserved before it could be used again. For this reason existing shared print networks have developed basic validation criteria to identify and reject copies in really bad shape, and indeed, this is in general what practitioners who work in circulating collections think of when then think about a book's condition. However, as we begin to think about the exigencies of shared print, I argue it's as important to ensure the copies we select for sharing are the most artifactually complete copies we can identify. In this sense "best copy" means mutually held title whose physical form is closest to the book in its original state. So for example, given three copies of a mutually held book where one copy has been rebound in library buckram, one is yet in its original publisher's binding, and one is yet in its original dust-jacket, the "best copy" would be copy #3.

Librarians who work in general collections are of course not used to thinking about books in this way. Traditionally, the physical or artifactual value of books is something to which our colleagues in special collections give attention and prioritize.

Former University of Pennsylvania Curator of Research Services Daniel Traister has written, "the root of the sense of the difference between general and special collections" has to do with preservation versus access: whereas in circulating collections access and the intellectual content of books is emphasized, in special collections preservation and artifactual value take precedence (para. III.2). And indeed, to the extent we "do" preservation in general collections, it is not to preserve the objects in which information is embedded but simply to ensure these objects last longer, even if this means destroying parts of the originals (e.g., rebinding books in buckram boards), or using surrogates (e.g., microfilm, or digital facsimiles) in their place. Additionally, in general collections we weed books based on criteria like circulation and condition, something that doesn't happen in special collections. In the preservation/access binary practitioners like Traister have posited, then, general collections librarians come down firmly on the side of access: for us a book's intellectual content or "intrinsic value" trumps its format or artifactual value, which is why there's such a thing as library bindings in the first place.

In recent years more than a few librarians and bibliographers, including Robert Bee, Michelle Cloonan, and G. Thomas Tanselle—and perhaps more (in)famously, bibliophiles like Nicholson Baker and Nicholas Basbanes—have challenged the general collections prioritization of access and intellectual content at the expense of original objects. The most well-known of these challenges is articulated in *Double-Fold: Libraries and the Assault on Paper* (New York: Random House, 2001), wherein Baker argues original objects should *never* be weeded, noting that more often than not surrogates—whether digital or otherwise—fail to reproduce original objects in accurate, adequate ways. Tanselle made similar arguments a few years before the publication of *Double-Fold*, suggesting that "when it is understood that access to physical evidence is an essential kind of access, and that books must *therefore be preserved in as many copies as possible*, the questions of ownership and care remain significant" (para. 5, emphasis added).

Of course, in addition to its not being entirely objective, *Double-Fold* was written more than a decade ago, during the earliest days of digitization, and this may be one of the reasons that today Baker's and Tanselle's positions seem not only extreme but also untenable. However, it's not the only reason: Roger C. Schonfeld and Ross Housewright note that Baker's argument to save (nearly) everything "cannot possibly be feasible when libraries hold hundreds of thousands of ill-used copies, far beyond the number required for access or preservation purposes" (p. 8). Instead, they suggest a strategic approach to deaccessioning print journals based on "a clear set of community preservation goals . . . that ensures preservation" (p. 9).

Although print monographs present different preservation issues than do print journals, like Schonfeld and Housewright I hope we can develop a thoughtful and rational preservation-centered strategy for the large-scale withdrawal of books likely to obtain in the wake of a shared print agreement. To this end I suggest adopting an artifact-focused view of preservation allied more closely with special, not general collections. For if one of the goals of shared print is to allow participating libraries to deaccession duplicate copies in order to free up space, then in a real sense we are creating scarcity where none existed before. In other words, whether shared copies will exist in a storage facility or not, in essence a shared print network will constitute a kind of new, special collection whose originals will have to be all things to future consortium member library researchers, including researchers interested in books as artifacts.

Developing the Survey Instrument and the Methods

In developing the survey instrument for my project I wanted the instrument to gather information about completeness of and damage to mutually held book copies in several consortial collections, but also about key artifactual elements of these items. I sought above all to keep my apparatus complex enough to capture significant artifactual, paratextual information, but simple enough for work study students to

deploy, and short enough to make analysis efficient and cost-effective. To help shape my questions I looked to some of the well-known published condition surveys undertaken in circulating collections at Yale, the University of Illinois, and Syracuse in the mid- and late-1980s; more recent surveys from the Universities of Kansas and Southern Mississippi; and a condition survey apparatus employed by the preservation unit at the University of California at Los Angeles. In part because my goals of hypothetical deselection of mutually held copies for shared print were different than the goals in these surveys (i.e., extrapolating conditions about entire collections, and prioritizing volumes in a single collection for preservation), without exception the survey instruments in these studies comprised far too many questions. However, the responses in the published studies informed my ultimate apparatus, which represents a kind of stripped-down version of these more complex surveys. (See <http://tinyurl.com/conditionsurvey> to view the instrument.)

The survey instrument I created is composed of five primary entries. Upon scanning an item's barcode the respondent is presented with three forced-response dropdown questions, two of which are followed by optional specifiers. Dropdown menus include nominal values describing the type of binding the item has, as well as 4-item Likert scales describing external and internal condition with options ranging from "very good" to "poor." The respondent is then asked whether to retain or discard the item, and a space for optional notes or comments concludes the instrument. Thirty-seven scripted responses to the five primary and four optional specifier entries would make data collection more efficient and data analysis easier, and also allow the respondent to make more explicit, objective claims about somewhat subjective Likert scale options.

The survey information was input directly into a web-based database created in Google Forms. I designed the form so that the barcode entry, forced-response dropdown questions, and up/down validation entry were required, obviating accidentally missed questions.

Additionally, for the sake of consistency I included explanatory notes for all but eight responses, which I deemed straightforward enough to stand alone.

Given concerns about the time (and cost) needed to undertake item-level condition analysis, it bears noting that leveraging Google Forms to design a survey instrument that fed directly into a web-based database, in addition to using barcodes as unique identifiers, made the process of data collection and analysis far easier and more efficient. Google Forms allowed for a simple, clean, and instructive data form. Scanning barcodes rather than inputting this information manually—or inputting another kind of unique identifiers like call number, title, author, or imprint information—saved inputting time. It also allowed me, post-survey, to draw out information about book copies from existing ILS item records and to manipulate this data for the purposes of comparing mutually held titles.

Before the survey could begin, a sampling method needed to be developed. I had significant help from USC Associate Dean for Collections John McDonald and SCELCL Program Manager Jason Price. From an existing dataset of OCLC numbers of holdings at SCELCL member libraries, McDonald derived a convenience sample of nearly 42,000 titles at my institution, Whittier College, that were published before 2010 and also held at two or more other SCELCL libraries. Next, Price, McDonald, and I met at the SCELCL offices to consider the feasibility of my Azusa Pacific University, BIOLA University, the Claremont Colleges, Loyola Marymount University, Mount Saint Mary's College, Pepperdine University, Whittier College, and the University of La Verne. See <http://tinyurl.com/kjos29w> for a map of the institutions.

It was determined that to generate statistically significant results, a final sample of approximately 4,000 items was necessary; and because the eight institutions hold these titles to varying degrees, items were sorted into categories based on the number of libraries in which they appear (3, 4, 5, 6, 7, 8). I strove to examine titles from each category in equal amounts, requiring the sample to include approximately 667 items per category.

This evenly distributed final sample was achieved by sorting the existing sample of available titles at the selected institutions by imprint date followed by call number, and then selecting every n th title in each category, where n was determined by dividing the total number of titles in each category by the number needed to result in the examination of 667 items. The sample was distributed thus:

Number of SCELCL libraries that own	Titles count	Copies count
8	87	696
7	94	658
6	112	672
5	134	670
4	168	672
3	223	669
Total	818	4037

Survey Implementation, Data Collection, and Manipulation

I began with an Excel file generated by Price with each title listed in 818 rows, and whose 11 columns included OCLC #, Author, Title, and then a listing of the eight SCELCL institutions; if a given title was held at a member library, this cell was marked with an "X." In order to isolate holdings at each library, I sorted the column that corresponded with the school I was visiting, deleted the remaining titles, and then sorted by call number. Next I printed these sheets out use as a checklist while I was implementing the survey.

Following a first survey conducted at my institution in late May, I visited the remaining seven SCELCL libraries between July 14 and July 27, a period at most places when books are not as likely to circulate. Prior to arriving at the libraries I contacted directors and staff to explain my project and arrange my visit, noting that for my purposes I would require a library cart for my laptop and scanner and the wireless password in order to access my Google Form, Worldcat, and local catalogs. With the exception of one institution—Loyola Marymount University, whose student workers pulled and shared titles prior to my

arrival—at each library I set my laptop and scanner on a borrowed cart, then proceeded to locate each item in the stacks, scan its barcode, examine the book and record the data in my form, then reshelfed the book before moving on to the next title. (See <http://tinyurl.com/nrcfk32> to see a few setup examples in situ.)

Following data collection, from the survey results spreadsheets I isolated the barcodes for the items I scanned at each institution. I then emailed these barcodes back to staff at each of the eight survey institutions, where systems librarians used review files to associate the correct author, title, and OCLC number with the barcodes, as well as the circulation data for these items, and then exported this information into a text file which they sent back to me. Next, I pasted this information into the survey results spreadsheets from Google Forms, color coded the data for each institution, and finally aggregated all the survey results in one spreadsheet. Arranging the data by OCLC number resulted in groupings of mutually held copies.

Results

In total I examined 3,429 book copies, spending two days at six libraries and one day at two libraries, where the average time to find and examine mutually held book copies was 90 seconds, i.e., around 40 books per hour. The majority of book copies I was not able to verify (i.e., locate in the stacks) were checked out to patrons or, as in the case of Azusa Pacific University, in the midst of a relocation. In the coming months I intend to return to three institutions in order to examine those copies I did not verify in this initial investigation. After examining and recording the conditions of these 3,429 copies, I compared mutually held titles in my aggregate spreadsheet. Three findings are, I think, worth sharing:

First, I discovered that the vast majority of the copies I examined are in what we might call “good shape.” In other words, only 2% of all books I examined had external conditions I regarded as poor, and only 1% of all books I examined had poor internal conditions. In short, I determined

that 98% of all the books I examined reasonably could be candidates for use in a shared print repository. This was somewhat surprising, especially after reviewing condition survey literature from several decades ago which led me to believe that a far greater number of books would be damaged or brittle.

Second—and perhaps unsurprisingly—though not a strong one, there is a correlation between the frequency a copy circulates and the extent to which it is damaged.

Third, and I think more importantly, when I plotted total copies against those copies that had what I designated “paratextual value” (i.e., original dust-jackets, original paperback binding, or facsimile paperback binding), then grouped by “total copies,” a clear trend emerged: overall, 31% of the copies in the groupings have paratextual value, which possibly indicates that if a title exists in less than 3 copies, *any* deselection has the potential to remove artifactually valuable copies from the shared print collective.

(See <http://tinyurl.com/p4jt9pn>, <http://tinyurl.com/oe25rg>, and <http://tinyurl.com/ps6s8qd> for graphic representations of these findings.)

Conclusions

For the purposes of a shared print agreement, from a condition perspective rooted in the culture of general collections where access to information is paramount, the overwhelming majority of the books examined for this project were good enough to retain. However, data also indicated that just one-third of the books in each copy grouping had paratextual value in the form of elements like original dust-jackets or publisher’s paperback bindings. This finding is significant from a more artifact-centered position that seeks to retain mutually held copies that have the most artifactual value, as it indicates not only that the majority of books in the study did not have paratextual elements, but also that random deselection is likely to result in the loss of artifactually significant copies.

References

- Schonfeld, R. and Housewright, R. (2009). What to withdraw? Print collections management in the wake of digitization. Retrieved from http://www.sr.ithaka.org/sites/default/files/reports/What_to_Withdraw_Print_Collections_Management_in_the_Wake_of_Digitization.pdf
- Tanselle, G. T. (1998) Texts and artifacts in the electronic era. *21stC*, 3.2. Retrieved from <http://www.columbia.edu/cu/21stC/issue-3.2/tanselle.html>
- Traister, D. (1998). Is there a future for special collections? And should there be?—A polemical essay. Retrieved from <https://www.english.upenn.edu/~traister/future.html>