Applying the Tiers of Assessment:

A Holistic and Systematic Approach to Assessing Library Collections

Madeline Kelly

George Mason University Libraries

4400 University Drive, MS 2FL

Fairfax, VA 22030

mkelly25@gmu.edu

(703) 993-2849

Keywords: Collections assessment, Academic libraries

Abstract

Collection assessment is a key component of collection development, budget allocation, and justification of library collections. Unfortunately, comprehensive collection assessment is daunting, subject to the weaknesses of individual tools and the overwhelming number of subject areas to assess. Few studies have attempted systematic assessment projects using multiple tools or methods, nor have many attempted to assess an entire collection subject-by-subject. This study implements an alternative to the single-tool model, combining multi-tool analysis with a systematic, subject-by-subject approach to the collection. The goal was to determine whether such a model of collections assessment was feasible in an academic library setting, providing usable data without overinvestment of manpower and resources. To this end, the method was tested in a pilot program at George Mason University (Mason), assessing three subjects at varying levels of depth. While there was concern that the methodology would prove too ambitious for full-scale implementation, the pilot yielded valuable, tangible results in a timely manner and provides a solid model for future assessment efforts at Mason and elsewhere.

Introduction

In this time of rapid technological, financial, and organizational change, libraries must constantly justify their collections. Each institution has to ask: Are our holdings worthwhile? What are our strengths? What are our weaknesses? These questions help librarians channel their efforts and their funding toward the resources that best suit the mission and needs of the institution. Collection assessment, the process by which one can answer these questions, is an essential component of ensuring that library expenditures are aligning properly with library goals. Yet, in a 2011 survey of 127 academic and special libraries, only 11% of respondents had

a formal process for collections assessment (Wilde & Level 2011). Systematic assessment, the objective alternative to gut feelings and anecdotal evidence, remains elusive for all but a few libraries.

George Mason University (Mason) is a large, relatively young public university in Fairfax, Virginia. Founded in 1957 as a satellite campus of the University of Virginia (and as an independent institution in 1972), Mason has grown rapidly in the last two decades and is positioned to continue expanding. In 2014, the George Mason University Board of Visitors approved a new ten-year strategic plan that includes the goal to "Achieve Carnegie Very High Research classification" (George Mason University 2014, p.24). The classification, which is based on the resources devoted by a university to research and development (R&D), is the highest level assigned to Doctorate-granting institutions (Carnegie Foundation 2014).

As Mason has expanded, the University Libraries have grown to keep pace with the University's institutional ambitions. The Libraries now include three distributed facilities in addition to the main research library. Collections have increased from 375,000 volumes in 1993 to more than 1.4 million volumes today, plus over 800 databases, 100,000 electronic journals, and 1.4 million e-books (Rein et al 1993; George Mason University Libraries 2013). The Libraries also house extensive archival materials within the Special Collections & Archives; the reference department, once staffed by just ten people, now includes over 15 faculty-rank librarians and a handful of classified staff. Moving forward, the Libraries hope to continue this growth, and have developed a unit strategic plan to complement Mason's broader efforts. The Libraries' plan includes two rounds of Balanced Scorecard assessment to benchmark collections and services and measure progress over two five-year periods.

From a collections perspective, the Balanced Scorecard will look at whether library collections reflect the needs of a very high research level institution. Measures of success will include an increase in the current collecting levels for select subjects, an increase in historic collection levels for the entire collection, and an increase in the Libraries' theoretical ARL ranking (based on ARL collection and budget metrics). Central to these efforts will be the assessment of current holdings and the identification of strengths, weaknesses, and opportunities for action. At least some of the Libraries' holdings (most likely select subject areas of significance to the University Strategic Plan) will need to be assessed systematically and repeatedly, at the beginning and end of each 5-year Balanced Scorecard cycle, to determine whether collection levels have actually improved.

Unfortunately, while the University Libraries are conscious of the need for collection assessment, several factors make assessment especially difficult at Mason. First, high turnover in the collection development department has stymied past efforts to establish and sustain assessment programs. Adding to this, many past efforts placed a heavy burden on subject specialists. These specialists, referred to as liaison librarians, were already expected to perform a wide range of job duties and simply did not have the time and energy to spend on a lengthy, subject-specific assessment process. Not only did the added liaison responsibility cause friction in the moment, it soured many library staff on the very idea of collection assessment, risking future efforts. Finally, the size of Mason's collections makes assessment an inherently complex and time-consuming endeavor. Even the simplest assessment could take years to complete thoroughly. Combined, these factors have prevented major progress on collection assessment, and any successful assessment program implemented at Mason must address these issues. The

goal of this study is to determine whether systematic assessment of library collections can

succeed in overcoming these hurdles by taking a flexible, systematic approach.

The Assessment Literature

There is no shortage of published collection assessment articles, from case studies to

literature reviews. The purpose of each assessment project varies, but common goals relate to

acquisitions, digitization, preservation, storage, and – particularly – weeding (Knievel et al

2006). What is consistent among many of these studies is that they are method-specific, applying

a single tool or strategy to the task of assessment. Many are also limited in that they assess only a

single subject area, or attempt to assess an entire collection with no attention to subject or

discipline. Examples of the single tool approach include Beile, Boote and Killingsworth's

citation analysis of education dissertations (2004); Monroe-Gulick and Currie's (2011) and

Spires' (2006) use of OCLC's WorldCat Collection Analysis; Beals and Gilmour's (2007)

experience with brief tests of zoology collections; and Kohn's (2013) application of usage data to

a whole collection. Example studies are innumerable, but most are method-specific and employ

only a single assessment tool or strategy.

The greatest weakness of this single-method approach to collection assessment is that it

hinges on the idiosyncrasies of the particular tool or strategy. This is problematic, given the well-

documented shortcomings of most assessment tools. Oke and Tyrrell summarize a few of the

challenges concisely (1999). For example, the Conspectus method (in which collections are

subdivided and examined meticulously), is prohibitively labor-intensive. It is also less

standardized than it would appear, as it depends on call number mapping, which rarely works as

objectively in practice as in theory (particularly in the face of inter- or multi-disciplinary

subjects). The Conspectus method is also one-dimensional, relying overly much on collection size and overlooking factors like use and quality. Checklists, which seem like an easier way to assess collections quickly and without call number mapping, are also problematic. The most authoritative list is still, to some extent, arbitrary. Worse, lists become outdated, encourage conformity, and – like the Conspectus – focus on quantity. Even seemingly objective measures like circulation statistics are one-dimensional, counting the use of resources but overlooking weaknesses in the collection, non-circulating materials, and the quality of available resources.

Wiemers et al summarize major collection assessment tools and their shortcomings further (1984). Accreditation standards, a convenient external measure, can be vague and arbitrary. Lists may be inappropriate to the local mission, as well as being time-consuming to check and frequently outdated. Citation analysis cannot easily be applied to broad fields with many sub-disciplines, since too many sources would be needed to sample the full range of core topics, and it does not address core materials that are used for background research without being cited. User studies, appealing because they reflect actual behavior, are flawed since not only may users be inexpert searchers, but their past preferences may not necessarily predict future use. Even major, commercial assessment tools have their shortcomings: OCLC's now-retired WorldCat Collection Analysis (replaced by a similar application, WorldShare Analytics Collection Evaluation) was based on notoriously inconsistent WorldCat data (Beals & Gilmour 2007) and had a tendency to over-report uniqueness and under-report overlap (Orcutt & Powell 2006). In short, none of the tools and methods covered in the literature is perfect. Collection assessment cannot depend entirely on any one of the available options.

The Cumulative or Multidimensional Approach

After reflecting on the strengths and weaknesses of each individual tool described in the literature, as well as the limited usefulness of assessing only a small section of the collection, Mason's solution was to apply multiple tools and measures in a holistic, well-rounded approach to assessment. Unlike single-tool assessments, the holistic approach uses each tool as just a small piece of the greater picture, pulling together a variety of quantitative, qualitative, internal, and external measures. This approach is similar to Oke and Tyrrell's "cumulative approach to collection evaluation" (1999), in which the collection is broken into subject-specific chunks and assessed systematically using a portfolio of tools. Oke and Tyrrell's cumulative approach included the use of peer comparisons, list-checking, and circulation statistics. After assessing each subject area, the researchers could then rate collections based on qualitative and quantitative results. The principles behind the cumulative model are echoed in Bodi and Maier-O'Shea's principles of assessment, which suggest:

> "-breaking down assessment by subject or smaller sub-topics when necessary
> -blending a variety of assessment tools appropriate to the discipline; and
> -matching print and electronic collections to departmental learning outcomes through
>   communication with faculty members." (2005, p.146)

These principles encourage a systematic but holistic approach. Further, unlike the many systematic assessment programs aimed purely at weeding, Bodi and Maier-O'Shea set out to assess the collections at North Park University for more open-ended, strengths-oriented purposes, which echo goals like those at Mason.[1] Thus, their principles are more broadly applicable than those intended only for weeding and adaptable to a variety of assessment aims.

---

[1] A similar approach was adopted in Finland by the Finnish Collection Map Consortium to map the strengths of libraries nationwide (Hyödynmaa, Ahlholm-Kannisto and Nurminen 2010). Various assessment tools were applied, though the need for standardization limited the use of qualitative data and the ability of individual libraries to locally customize assessment.

In 2008, Borin and Yi condensed many of these concepts, proposing a "new dimensional framework" for evaluating collections. Like Bodi and Maier-O'Shea, they urge the use of multiple assessment tools to create a clearer picture of the collection. They also emphasize the use of various perspectives, including usage-based perspectives and standards-based perspectives. Where Bodi and Maier-O'Shea share case study experience, Borin and Yi build a flexible model that can be tailored to local need. Combined, these studies in multidimensional collection assessment make a compelling case for a more comprehensive approach.

Mason's collection assessment program grew out of these theories and models, based on the basic idea that multiple perspectives are essential to good assessment. Thus, Mason's approach is both holistic and piecemeal, addressing each subject area in turn and assessing materials with a portfolio of tools. Data collection and analysis follow standardized procedures and templates that can be molded to accommodate the idiosyncrasies of each discipline. Each subject collection is assessed using internal and external measures, including peer comparisons, and then assigned a rating based on the entire range of results. (The list of tools used is described in detail in the methodology, below, and in Appendix A.) Assessment at Mason is the practical implementation of previous models with a comprehensive scope and systematic approach.

While the holistic, subject-by-subject approach is ambitious, taking far longer than a single-tool or single-subject study, the results should be richer and more useful than smaller studies. Mason's approach will allow library staff to move through subject areas systematically, and while some subject areas will have to wait a year or more for their turn at assessment, each actual assessment will take just a few months and will yield usable data at the end of that period. As more subjects are assessed, the library will enjoy a growing pool of results and an increasingly meaningful overview of the entire collection.

Other benefits of the approach include the fact that it is broad, drawing from qualitative, quantitative, internal, and external measures (see Table 1). It is also flexible and customizable, allowing for variations from subject area to subject area. Smaller subjects, or subjects of lesser importance to the institutional mission, can be assessed at a lower intensity than large subject areas supporting PhD programs of vital importance to the University. In addition, the approach is cyclical and incremental: the assessment process is applied to each subject area in turn, cycling through the steps and working piece by piece through the entire collection. Procedures and templates can be revised and updated each time a new subject area is assessed, providing a built-in opportunity for evaluating assessment efforts.

Table 1. Collection assessment methods

Use- or User-Based	Collection-Based
Quantitative	*Quantitative*
Interlibrary loan statistics Citation analysis E-Journal & database usage	Bowker Book Analysis System Bibliographies OCLC Collection Evaluation Environmental scan of e-resources Brief test of collection level
Qualitative	*Qualitative*
Survey of faculty & students Publishing Outlets	Accreditation guidelines
Based on Matthews' chart of evaluation methods (2007, p.112).	

Finally, the Mason model of collection assessment departs from convention in one key way: it is a centralized effort designed to rely as much as possible on a single individual. Wilde and Level suggest that assessment "requires huge amounts of interdepartmental collaboration as well as the expertise of library administrators, professional librarians, and support staff" (2011, p.223). Unfortunately, Mason's assessment projects have faltered in the past precisely because

they involved such huge amounts of work by so many individuals. Any new efforts must defy

the intensely collaborative model, placing the core of responsibility with a collection

development staff member, using outside expertise only as necessary. After the planning stages,

which inevitably involve the collaboration and input of many stakeholders, the program can run

nearly on its own under the responsibility of a single position.

Mason's Pilot Program: Overview & Methodology

After several months of development, a pilot version of Mason's collection assessment

program was launched in October 2013. In order to test the program's effectiveness across a

range of disciplines, three subjects were selected for the pilot: linguistics, arts management, and

forensic science. These subjects had the added benefit of corresponding to relatively small

programs or departments, which helped minimize the challenges of implementing the pilot

program. The assessments were conducted serially, allowing for adjustments to the methodology

along the way. The resulting procedure, developed and honed over the course of eight months, is

explained in greater detail below.

Each assessment began with a consultation between the appropriate liaison librarian and

the staff member tasked with assessment. Liaisons completed a short survey on their subject area

and any specific assessment concerns (see Appendix A). Liaisons supplied assessment staff with

two lists of peer libraries (one current, one aspirational) as well as detailed call number ranges,

lists of top journals and databases, assessment goals, and other potentially useful information.

Liaisons were granted some flexibility in the extent of information they provided, as well as the

extent to which they participated in the assessment itself. Finally, liaisons selected the level of

detail they wanted from assessment; options ranged from basic assessment (Tier I) to

intermediate (Tier II) to thorough (Tier III). Liaisons were granted a few weeks each to complete

the survey, minimizing inconvenience. Armed with the liaison's knowledge, perspective, and preferences, assessment staff could then proceed with gathering the data.

Data gathering followed a standardized procedure, using a portfolio of tools and measures organized around the Tiers of Assessment (see Appendix B). Lower Tiers included tools that are easy to implement, making it feasible to apply a base level of assessment to all subject areas. Higher Tiers included all the tools used in lower Tiers, with the addition of more nuanced and difficult-to-implement measures like user surveys and in-depth citation analysis. The increasing complexity of the tools across the Tiers is intended to yield increasing levels of depth and detail in the results. Arts management and forensic science were assessed at Tier II; linguistics at Tier III. For basic assessments (Tier I) the following tools and measures were used:

- Bowker Book Analysis System (BBAS). BBAS, an online subscription service, was used to compare Mason's holdings to the Choice list *Resources for College Libraries*, a list of undergraduate core resources.

- List-checking against bibliographies. Bibliographies were gathered online from reputable sources (including professional associations and peer institutions), then checked against local holdings. At least 400 titles, including journals, were gathered for each subject area.

- OCLC Collection Evaluation. Collection Evaluation, a subscription service, was used to compare Mason holdings to peer libraries' in terms of age, language, and uniqueness. Collection Evaluation also allowed Mason to compare holdings with Choice's *Outstanding Academic Titles*.

- List-checking against peer e-resources. E-resources (excluding journals) listed on peer institutions' subject guides were compared to Mason holdings and to each other.

- Brief test of collection strength. White's brief test, published in 1995, was used to determine the approximate level of the collection. Items were sampled from the bibliographies (above), ranked according to WorldCat holdings, and compared to Mason's collection. Item rankings and collection levels were assigned based on Lesniaski's modified version of White's original methodology (2004).

For intermediate assessments (Tier II), the following tools and measures were added:

- ILL statistics. Two years of ILL data were reviewed to measure the volume of loans and requests, as well as patron demographics. Data included separate sections for local consortium lending in addition to broader ILL numbers.

- Citation analysis of core journals to determine researcher preference in terms of age and format of materials. Twelve articles were selected, two each from six representative, subject-specific journals. Citations were sampled from each article (for a total of 200-250 citations), and format and age were recorded for each citation. Results were used to calculate an ideal serial/monograph ratio for each subject (shown in Table 3 as % Serials). Subject librarians were also supplied with lists of most-cited resources and the age distribution of these materials.

- Usage data for e-journals and databases. Core databases and journals suggested by the liaison librarian were checked for usage. Data included full-text downloads (for journals) and searches (for databases).

- Faculty and graduate student surveys. Two survey instruments were developed and administered electronically to willing faculty and students in each program or department. See Appendix C for the faculty survey instrument. (The student survey instrument parallels the faculty survey closely and has not been included.)

Thorough assessments (Tier III) included:

- Citation analysis of monograph resources. Eight core titles were selected from the brief

  test list (above). Titles were selected to represent different levels as well as all major

  topics or sub-disciplines within the overarching subject. Citations were sampled (200 in

  total) and processed as with journal citations, above. For thorough assessments,

  monograph citations were also combined with serials citations for a cumulative average

  of researchers' preferences.

- Accreditation guidelines, where applicable. While none of the test subjects had formal

  accreditation requirements, future assessment efforts will include any explicit standards

  provided by outside governing bodies.

As the data was collected, it was saved based on established file-name conventions within

a consistent folder structure; each assessment comprised a standardized set of files. The main

results of each subject assessment were compiled into a running document called the Progress

Checklist, based on a template to ensure that data was summarized consistently across all

subjects. Procedures for data collection were also documented in an Assessment Program

Manual, which underwent multiple revisions throughout the pilot program. The Manual, by

detailing the process, ensures the project will have continuity independent of staff changes.

Once all the data was gathered for a given subject, assessment staff performed any

necessary analysis and data manipulation, and then compiled key findings into a standardized

report. The reports were designed to be concise (approximately eight pages for a Tier III

assessment) and include brief overviews of the subject, its corresponding academic department,

and the assessment findings. The report also includes a list of supporting documents (i.e. raw

data files and other in-depth results) available to readers upon request. Reports were then

distributed in draft format to the Head of Collection Development and the relevant liaison

librarian for review. Based on the findings, assessment stakeholders collaborated to assign the

subject collection a rank based on the WLN/RLG Collection Level Indicators (see Table 2), and

to make recommendations for future collection development efforts.

Table 2. Summary of WLN/RLG Collection Level Indicators

1a Minimal Level, Uneven Coverage	Few selections; unsystematic.
1b Minimal Level, Even Coverage	Few selections; covers basic authors, core works, and a range of ideas.
2a Basic Information, Introductory	Introduces and defines topic; basic reference works and historical descriptions; general works on major topics. Some major periodicals.
2b Basic Information, Advanced	2a plus greater depth and more topics; editions of key works; supports needs of highly educated public or community college students.
3a Basic Instructional Support	Adequate for teaching and maintaining knowledge; key primary and secondary texts; basic serials and indexes; fundamental reference tools; supports lower undergraduates.
3b Intermediate Instructional Support	3a plus a broad range of basic works; retrospective materials; key journals on primary topics; some journals on secondary topics; supports advanced undergraduate work.
3c Advanced Instructional Support	3b plus ample works and journals on primary and secondary topics; substantial number of works by secondary figures; in-depth works on research, techniques, evaluation; supports graduate work and advanced independent study.
4 Research	Most published materials needed for dissertations or published research; all important reference works; wide selection of specialized monographs; extensive collection of journals; major abstracting and indexing services; foreign language materials; older materials.
5 Comprehensive	All significant works of recorded knowledge in all languages; exhaustive.

When the assessment team agreed on a final version of the report, it was sent to library

administration. Liaisons were issued a clean copy of the report along with a ZIP file containing

all associated assessment files and a guide for how to interpret them. This process was repeated

for all three test subjects over the course of eight months.

Pilot Program Challenges

Upon completion of the pilot program, it became clear that there were several challenges to this method of collection assessment. First, it was often not possible to achieve an appropriate level of granularity for the subject being assessed. Many subjects, like arts management, are interdisciplinary and hard to isolate as call number ranges. Others, like forensic science, are so specialized as to evade many of the commercial evaluation tools. While assessment staff strived to gather only the most relevant data, it was often impossible to separate the chaff from the wheat; similarly, future subject assessments may overlap with one another due to classification challenges. While the wide range of tools used mitigates some of this "fuzzy" data, accurate collection assessment for interdisciplinary or specialty subjects remains a challenge. Any assessment of those collections will have to be tempered by consultations with subject specialists and the results analyzed critically.

Beyond these subject-specific challenges, the tools and measures themselves each had strengths and weaknesses that influenced the quality of the data gathered. BBAS, for instance, could not account for Mason's e-book holdings, so all results had to be double-checked against the Mason catalog manually. There were also concerns among liaison librarians about the quality of the *Resources for College Libraries* list and the classification scheme used by RCL. OCLC's Collection Evaluation tool was similarly problematic: the OCLC classification scheme was only so precise. Further, due to the quality of WorldCat data, it likely over reported uniqueness (Orcutt & Powell 2006). Even the basic tools, like bibliographies and citation analysis, paint an inherently limited picture. Taken together, these imperfect tools may provide a reliable picture of a collection's strengths; taken alone, none is adequate to the task. The determined collection assessor must accept these flaws, mitigate them when possible, and move on.

Lastly, the inherent challenge of assessing collections is how to make meaningful comparisons and assign ratings in the absence of clear benchmarks. No single measure should be relied upon as the "final word" on the collection, and yet, without an objective, absolute measure, the final collection rating seems almost arbitrary. What makes for a 3C collection? What makes for a 2B collection? How should the various results be weighted? The WLN/RLG Collection Level Indicators – Mason's guiding authority on the final ratings – are only so specific, and with such an abundance of data it was often difficult to rule.

Pilot Program Successes

In spite of the pilot program's challenges, there were also many successes. First, the cyclical and incremental nature of the program proved worthwhile. Although the first subject assessed took several months, subsequent assessments moved more quickly, as procedures and templates were refined. In-depth assessment is inherently time-consuming, but there is hope that an iterative process can produce an efficient – if ambitious – model. Further, as each additional subject area was assessed, it became easier to benchmark and rate collections. With an increasing pool of data, decisions and recommendations become more rooted in context and less arbitrary (see Table 3). With comparison data, the process of assigning a collection level becomes more scientific. As assessment continues, the overall picture is likely to become clearer and richer, to the extent that Mason may someday be able to share some form of benchmarking guidelines with other institutions.

Table 3. Summary Results of Pilot Assessment

	Linguistics	**Arts Mgmt.**	**Forensic Science**
Level of Assessment	Tier III	Tier II	Tier II
Assigned Level	***3C***	***3C***	***3C***
Highest Degree Offered	PhD	MA	MS
Goal Level	4	3C	3C

Brief Test Level	3C	3C	4
Survey Level	3.83 (3C)	2.5 (2B-3B)	3.2 (3B)
% Serials (Ideal)	37%	49%	77%
% RCL Owned	90%	89%	89%
% Bibliographies Owned	75%	61%	62%
% Outstanding Ac. Titles	100%	79%	77%
% Peer DB Owned	77%	79%	62%
% Top Journals Owned (CA)	100%	100%	100%
% Top Journ. Owned (Survey)	73%	50%	53%
% Top Books Owned (CA)	43%	60%	n/a
% Top Conf/Series Own (CA)	28%	n/a	n/a
Peer Ranking (Size)	13 of 14	9 of 11	4 of 15
% of Holdings 1990 or Newer	77%	78%	83%
% Holdings 2000 or Newer	51%	53%	61%
% Non-English Materials	3.6%	3.2%	0.7%
Language Preference	60% English only	83% English only	100% English only
ILL Gradient	Net Lender	Net Borrower	Net Borrower
Prefer Print	55%	83%	57%
Prefer e-Book	30%	0%	22%

Another success of the model was its centralization. As had been hoped, a single staff member carried out most of the work, minimizing the burden on subject specialists. All procedures were documented thoroughly, anticipating inevitable staff changes and the need to train others on the assessment program. While liaison librarians were consulted intermittently throughout the process, their expertise was used strategically, and only as needed. In fact, the bulk of liaison involvement came at the end of the process, when the assessment yielded tangible results for their interpretation and use. Otherwise, assessment was a largely invisible force operating behind the scenes.

The program's greatest indicator of success, hinted at above, is the abundance of title lists, usage data, and peer comparisons now available to staff for collection development. Liaisons were able to purchase journal back files, monograph titles, and e-resources; ILL helped place Mason's collections in a local and global context; and peer comparisons helped liaisons

gauge the appropriate age and linguistic breakdown of the ideal collection. For liaisons new to the subject area, collection assessment provided insight into Mason's holdings and the needs of the liaisons' constituents. In short, the assessment pilot program yielded a wealth of information that is still being unpacked today. Abstract goals like "understanding Mason's collections," "facilitating liaison decisions," and "advancing the Libraries" were achieved with tangible, usable results.

Conclusions

Since the completion of the pilot, the holistic assessment methodology has been formally approved for implementation as a full-scale collection assessment program at Mason. Select subjects will be assessed multiple times throughout the 10-year Balanced Scorecard exercise as a way to monitor progress. Other subjects will also be assessed as part of a broader push to assess all of Mason's library holdings over the next several years. In spite of concerns that the methodology would be too time-consuming to be feasible, it has in fact proven fruitful. Further, the pilot program allowed librarians at Mason to tweak the methodology as needed before full-scale use. Some measures, such as circulation data and collection growth figures, were too challenging to gather and were discarded from the final program. Others, like the student and faculty surveys, proved more valuable than expected. As the program moves forward, other adjustments may need to be made to the portfolio of tools used and the scope of subjects to be assessed. A slightly abbreviated version of the methodology may be more appropriate in certain cases, especially for a large university hoping to complete a full-scale, systematic collections assessment in a scant few years. Fortunately, the structure of the approach allows for customization and should be simple to adjust.

Outside of Mason, the cumulative model tested here can and should be explored by other libraries as a way to assess collections thoroughly, flexibly, and without undue burden on subject specialists. The holistic approach allows each subject collection to be viewed from many angles, while gradually developing a picture of the broader collection as a whole. Best of all, the tools used are complementary, and patterns emerge as results reinforce or contradict one another. Whether it is applied to all subjects or just a few, or whether the tool kit consists of four tools or fourteen, the cumulative approach produces valuable results. Though not for the faint of heart, it is an effective, if ambitious, strategy for collection assessment, and one that might benefit university libraries of all aims and sizes.

References

Beals, J.B. & Gilmour, R. (2007). Assessing collections using brief tests and WorldCat

Collection Analysis. *Collection Building*, 26(4), 104-107.

Beile, P.M., Boote, D.N., & Killingsworth, E.K. (2004). A microscope or a mirror?: A question

of study validity regarding the use of dissertation citation analysis for evaluating research

collections. *Journal of Academic Librarianship*, 30(5), 347-353.

Bodi, S. & Maier-O'Shea, K. (2005). Library of Babel: Making sense of collection management

in a postmodern world. *Journal of Academic Librarianship*, 31(2), 143-150.

Borin, J. & Yi, H. (2008). Indicators for collection evaluation: A new dimensional framework.

Collection Building, 27(4), 136-143.

Carnegie Foundation. (2014). "Methodology." *Carnegie Classification*. Accessed 6 June 20114.

Retrieved from http://classifications.carnegiefoundation.org/methodology/basic.php .

George Mason University. (2014). *2014-2015 Strategic Plan.* Retrieved from

http://strategicplan.gmu.edu/wp-

content/uploads/2013/07/GMU_Strategic_Plan_Web.pdf.

George Mason University Libraries. (2013). "Facts & Figures." *About the Libraries.* Retrieved

from http://library.gmu.edu/about/facts.

Hyödynmaa, M., Ahlholm-Kannisto, A., & Nurminen, H. (2010). How to evaluate library

collections: A case study of collection mapping. *Collection Building*, 29(2), 43-49.

Knievel, J.E., Wicht, H. & Connaway, L.S. (2006). Use of circulation statistics and interlibrary

loan data in collection management. *College & Research Libraries*, 67(1), 35-49.

Kohn, K.C. (2013). Usage-based collection evaluation with a curricular focus. *College &*

Research Libraries, 74(1), 85-97.

Lesniaski, D. (2004). Evaluation collections: A discussion and extension of 'Brief Tests of

   Collection Strength.' *College & Undergraduate Libraries*, 11(1), 11-24.

Matthews, J.R. (2007). *Evaluation and measurement of library services.* Westport, CT: Libraries

   Unlimited.

Monroe-Gulick, A. & Currie, L. (2011). Using the WorldCat Collection Analysis tool:

   Experiences from the University of Kansas Libraries. *Collection Management*, 36, 203-

   216.

Oke, G. & Tyrrell, B. (1999). Cumulative approach to collection evaluation. In *Adelaide 98,*

   Pathways to Knowledge: Australian Library and Information Association 5th Biennial

   Conference and Exhibition, 25-28 October 1998, Adelaide Convention Centre:

   Conference Proceedings.

Orcutt, D. & Powell, T. (2006). Reflections on the OCLC WorldCat Collection Analysis Tool:

   We still need the next step. *Against the Grain*, 18(5), 10.

Rein, L.O., Hurley, F.P., Walsh, J.C., & Wu, A.C. (1993). Formula-based subject allocation: A

   practical approach. *Collection Management*, 17(4), 25-48.

Spires, T. (2006). Using OCLC's WorldCat Collection Analysis to evaluate peer institutions.

   Illinois Libraries, 86(2), 11-19.

White, H.D. (1995). *Brief tests of collection strength*. Westport, CT: Greenwood.

Wiemers, E., Baldwin, C.A., Kautz, B., Albrecht, J., and Lomker, L.H. (1984). Collection

   evaluation: A practical guide to the literature. *Library Acquisitions: Practice and Theory*,

   8, 65-76.

Wilde, M. & Level, A. (2011). How to drink from a fire hose without drowning: Collection

   assessment in a numbers-driven environment. *Collection Management*, 36(4), 217-236.

Appendices

Appendix A: Liaison Survey

[LIAISON], [FUND]

Please answer the following questions to help us more effectively assess your subject collection.

1. How extensive an assessment would you like for this subject?
   __ GENERAL (lowest level; ideal for small or lower-priority programs)
   __ MODERATE (mid-level; ideal for mid-sized or up-and-coming programs)
   __ THOROUGH (highest level; ideal for large programs, high-priority programs, or
      programs of special interest to the University)

2. Have you recently, or as part of your routine, assessed your collection in any way? If so,
   how?

3. Is there any aspect of your collection that you would particularly like measured?

4. What unique challenges – if any – will your subject area pose to the assessment process?
   (E.g. unusual formats, multiple languages, interdisciplinary focus, etc.)

5. Is your collection development profile up-to-date, including curricula and LC call numbers?
   __ NO   __ YES
   *If NO, please contact CDP to update your collection development profile.

6. Identify 3-7 current peer libraries in your subject area. Please provide a 1-2 sentence
   explanation of why you have selected each peer or a paragraph describing the group as a
   whole.

7. Identify 3-5 aspirational peer libraries in your subject area. Please provide a 1-2 sentence
   explanation of why you have selected each peer or a paragraph describing the group as a
   whole.

8. Please indicate your goal collection level based on the following:
   __ 1a Minimal Level, Uneven Coverage. Few selections; unsystematic.
   __ 1b Minimal Level, Even Coverage. Few selections; covers basic authors, core works, and
      a range of ideas.
   __ 2a Basic Information, Introductory. Introduces and defines topic; basic reference works
      and historical descriptions; general works on major topics. Some major periodicals.
   __ 2b Basic Information, Advanced. 2a plus greater depth and more topics; editions of key
      works; supports needs of highly educated public or community college students.
   __ 3a Basic Instructional Support. Adequate for teaching and maintaining knowledge; key
      primary and secondary texts; basic serials and indexes; fundamental reference tools;
      supports lower undergraduates.

__ 3b Intermediate Instructional Support. 3a plus a broad range of basic works; retrospective materials; key journals on primary topics; some journals on secondary topics; supports advanced undergraduate work.

__ 3c Advanced Instructional Support. 3b plus ample works and journals on primary and secondary topics; substantial number of works by secondary figures; in-depth works on research, techniques, evaluation; supports graduate work and advanced independent study.

__ 4 Research. Most published materials needed for dissertations or published research; all important reference works; wide selection of specialized monographs; extensive collection of journals; major abstracting and indexing services; foreign language materials; older materials.

__ 5 Comprehensive. All significant works of recorded knowledge in all languages; exhaustive.

For more information on these collection levels, visit http://goo.gl/B0xjXc .

9. How involved would you like to be in the assessment process? (Check all that apply.)
    __ I would like to suggest bibliographies for the list-based portions of the assessment
    __ I would like to suggest key journals and monographs for citation analysis
    __ I would like to suggest key journals and databases for usage statistics
    __ I would like to involve department faculty in suggesting resources for study
    __ I would like to help analyze and summarize the data
    __ I would like to participate in writing the final report
    __ I would prefer not to be involved until a draft of the final report has been written
    __ Other/Comments:

10. Would you be willing to distribute a survey to faculty affiliated with your program?
    __ NO   __ YES

11. Would you be willing to distribute a survey to graduate students affiliated with your program?
    __ NO   __ YES

12. Please list your faculty liaison and/or any faculty members who might be willing to participate in the assessment process (beyond survey completion):

13. What outcomes do you hope for from the assessment process?

14. Please include any other comments that come to mind regarding assessment:

Appendix B: Tiers of Assessment

All Tiers (General)
- Bowker Book Analysis System (comparison to core titles from Resources for College Libraries)
- Basic bibliographies (comparison to reputable lists of core titles)

- OCLC Collection Evaluation (comparison to peer libraries and Outstanding Academic Titles)
- Environmental Scan of e-Resources (among peers)
- Brief Test (succinct method for estimating collection level)

Tier Two (Moderate) and Above
Add:
- Interlibrary Loan Statistics (items loaned and borrowed)
- Analysis of citations from key journals to determine preferred format and age of materials (journals identified using bibliographies and liaison input; selected based on ranking and scope)
- E-Journal Usage
- Database Usage (of subject-specific databases)
- Survey of faculty (standard survey instrument across all subjects)
- Survey of graduate students (standard survey instrument across all subjects)

Tier Three (Thorough)
Add:
- Analysis of citations from monograph sources to determine preferred format and age of materials (monographs identified using bibliographies and liaison input; selected to represent the full scope of the subject)
- Combined analysis of citations from journals and monographs (to determine the average preferred format and age of materials)
- Accreditation guidelines (where applicable)

*Tiers to be reevaluated after each round of assessment.

Appendix C: Faculty Survey
[PROGRAM(S)]

DEMOGRAPHICS (1 of 4)
1. For what reason(s) do you use the Mason libraries? Mark all that apply; indicate your primary reason with an asterisk.
   __ Support teaching   __ Personal (non-research)
   __ Research   __ Other:

2. With what department(s) or program(s) are you affiliated?

3. What is your title/faculty rank?
   __ Instructor   __ University Professor
   __ Adjunct   __ Distinguished Service Professor
   __ Assistant Professor   __ Emeritus Faculty
   __ Associate Professor   __ Other:
   __ Professor

4. What level do you teach?

__ Undergraduate   __ Post-Graduate
__ Graduate

5. Within your subject, what are your research interests and/or areas of expertise?

6. How long have you taught at George Mason?

FREQUENCY OF USE (2 of 4)
7. On average, how often do you visit one of Mason's <u>physical</u> libraries?
__ Every day   __ A few times per semester
__ A few times per week   __ A few times per year
__ A few times per month   __ Twice a year or less

8. Which Mason libraries do you use? Mark all that apply; indicate your preferred location with an asterisk.
__ Fenwick Library (Fairfax)   __ Arlington Campus Library
__ Johnson Center Library (Fairfax)   (Arlington)
__ Mercer Library (Prince William)   __ Law Library (Arlington)

9. On average, how often do you access online resources provided by the University Libraries (including but not limited to: databases, e-journals, e-books, datasets, streaming content, etc.)?
__ Every day   __ A few times per semester
__ A few times per week   __ A few times per year
__ A few times per month   __ Twice a year or less

10. Do you use other libraries outside of Mason? If yes: Which one(s), and why?
__ No   __ Yes:

MATERIALS USED (3 of 4)
11. Within your subject, what online resources do you use (generally and/or specifically)? These can include (but are not limited to) databases, e-journals, e-books, Infoguides, subject portals, datasets, streaming content, etc. as well as specific titles/products.

12. Within your subject, what print resources do you use (generally and/or specifically)? These can include types of materials as well as specific titles/products.

13. Please list up to five journals that you consider to be core resources for your subject (print journals or electronic):

14. For your professional work, do you consult materials in any language(s) other than English? If yes, what language(s)?
__ No   __ Yes:

15. For your professional work, do you tend to prefer e-books or print books? Why?

GENERAL RESPONSES (4 of 4)

16. Do you have a good sense of the University Libraries collection(s) for your subject? Feel free to explain.
    __ No   __ Yes:

17. Can you identify weaknesses in the University Libraries collection(s) for your subject? (Please include weaknesses at all levels, from core essential materials to specialized research materials.)
    __ No   __ Yes:

18. Can you identify strengths in the University Libraries collection(s) for your subject? (Please include strengths at all levels, from core essential materials to specialized research materials.)
    __ No   __ Yes:

19. Please rate the University Libraries collection(s) for your subject on a scale of 1 to 5, based on the following guidelines:
    __ 1. Minimal Level: Includes basic works.
    __ 2. Basic Information Level: Includes up-to-date general materials that serve to introduce and define a subject; not sufficiently intensive to support courses or independent study in the subject area.
    __ 3. Instructional Support Level: Adequate to support undergraduate and most graduate instruction as well as independent study; sufficient to maintain knowledge of a subject for general purposes, but insufficient for advanced research.
    __ 4. Research Level: Includes major source materials for dissertations and independent research; includes all major reference works, a wide range of publication dates, and specialized monographs, journals, and databases.
    __ 5. Comprehensive Level: Includes all significant works, in all applicable languages, relevant to the field; exhaustive.

20. What way(s) – if any – do you feel that the library collections have impacted your scholarship?

Thank you for your participation in our survey. If you are interested in a follow-up, please include your name and contact information – as well as any questions, concerns, or comments you may have – below: