Examining the “Culture of Assessment” in Public Libraries

Allison Ferguson¹

¹ School of Library & Information Studies, University of Alberta, Email: apfergus@ualberta.ca

Abstract

The introduction of the New Public Management (NPM) approach to public libraries globally brought with it a quality management paradigm typical of that used in for-profit corporations, characterized by an audit-intensive environment. This led to a desire for a “culture of assessment” to be the preferred standard within public libraries, largely in response to the threat of budget reductions and library closures. This review examines literature from the last two decades on the impact of this strategy, including themes such as the apparent consensus around library value (both in terms of qualitative facets and quantitative economic measures such as return on investment) and traditional and more recent approaches to measuring that value in both research and professional practice publications. Issues identified in the literature with the quantitative emphasis and threat narrative of the culture of assessment are discussed, as well as studies which examine the direct impact of NPM’s neoliberal ideology on public libraries.

Keywords: public library, public libraries, evaluation, outcome assessment

In 2004, Lakos & Phipps wrote the seminal work on a phenomenon rapidly gaining traction within libraries: the “culture of assessment”. According to their definition, a culture of assessment is “an organizational environment in which decisions are based on facts, research, and analysis…services are planned and delivered in ways that maximize positive outcomes for customers and stakeholders…staff care to know what results they produce and how those results relate to customers’ expectations” (p. 352).
This outlook is consistent with the New Public Management (NPM) approach, which emphasizes, among other things, an audit-intensive environment (Düren et al., 2017; Irwin & St-Pierre, 2014; Steane et al., 2015). Lakos & Phipps (2004) also state that “libraries must transform themselves into organizations that value quality and quality management” (p. 347). This context has led to an adoption of quality management approaches common in for-profit corporations, including a focus on such tools as key performance indicators (and associated targets), cost-benefit analyses, and quality scorecards (de la Mano & Creaser, 2016; Reid, 2020). Similarly, to their use in the private sector, these outputs are designed to provide primarily quantitative assessment of library “performance” to the group Halpin et al. (2015) refer to as the “power people”, decision makers who ultimately impact library funding as opposed to patrons or library staff (p. 36).

This literature review examines the current impact of the culture of assessment specifically within public libraries, via exploration of recent published works (both academic and professional practice-oriented). It takes a broader scope than other literature reviews focused on library value (Aabø, 2009; Imholz & Arns, 2007; Matthews, 2011; Sørensen, 2020; Stenstrom et al., 2019; Vårheim et al., 2019) in that it seeks to examine not only the practices used for assessment in public libraries, but the implications of those practices. Themes that arose in the literature will be examined in depth, including the presence of a consensus on how the public library provides value, benefits, and shortcomings of common evaluation methods (particularly in the fields of economic valuation and outcome assessment), and the results of the corporate quality approach as implemented within the public library context.

Methods

Significant effort within both academic and professional organizations has been made in the field of library evaluation, for all varieties of libraries. The present review is focused on public libraries, as assessment in academic libraries is often driven by university-wide assessment strategies and has a consequently narrower focus on student and faculty requirements (Oakleaf, 2010). While certain assessment techniques and methodologies will certainly overlap, the purpose of this review is primarily to
examine how the use of these methodologies are currently affecting public libraries in the context of the culture of assessment. As such, included references were subject to the following criteria:

- Resources (articles/books) describing methods and studies must have the public library context as the primary focus.
- Resources must be peer-reviewed, or published by a recognized professional organization (e.g., American Library Association) from 2001 to 2021, with emphasis given to studies published between 2016 and 2021.
- Resources must be in English (no geographic limitations).
- Review articles and meta-analyses were included if other criteria were met (reference lists were checked for substantial overlap before conclusions of these articles were presented as independent results).

To initially identify eligible resources, the above criteria were applied to subject search for ("public library" OR "public libraries") AND ("assessment" OR "evaluation") within the EBSCO Library and Information Science Source database (69 results) and the ProQuest Library & Information Science Abstracts (LISA) database (229 results). Review of titles and abstracts to isolate studies addressing the chosen topic, plus additional studies added based on references within resulted in a final sample of 47 studies.

**Analysis and Commentary**

**Value of Public Libraries: Central Ideas and Common Facets**

An interesting theme that emerged from reviewing the selected studies is the general consensus around the areas in which libraries impart value to the public (Halpin et al., 2015; Matthews, 2011; McMenemy, 2007; Sørensen, 2020; Stenstrom et al., 2019; Vårheim et al., 2019), even amongst municipal decision makers who may or may not be public library users themselves (Gazo, 2011). Stenstrom et al. (2019), in a review of over 130 resources from 1998-2018, identified three “umbrella” areas of support provided by libraries: support for personal advancement (e.g., knowledge, emotional wellness), support for vulnerable populations (e.g., those experiencing homelessness, immigrants and literacy learners), and support for community development (e.g., social infrastructure and community resilience). A systemic review by Vårheim et al. (2019), on
current literature of the library as a public sphere institution, also calls out community
development and explicitly identifies the service areas the public library offers to create
these supports: community focused services such as meeting places and specific
programs for marginalized populations, knowledge organization, provision of new tools
and services, institutional structures and practices, and sources of funding.

While there is agreement within the literature on the ways in which libraries
provide value, there is less concurrence on the methods by which that value is
measured. Are the crucial measures of library value strictly in economic value, as in
studies which address the “return on investment” provided by public libraries (Aabø,
2009; Imholz & Arns, 2007)? Where these cumulative measures do not provide
sufficient detail, how should libraries assess outcomes and impacts on a program or
service level (Lyons, 2012)? In the following sections, a review of recent literature on
both traditional and current approaches is presented.

**Traditional Methods of Library Assessment**

**Library Valuation and the Rise of Corporate Metrics**

Beginning in the early 2000s, in response to the same concerns expressed in
Lakos & Phipps (2004), substantial effort was made in the literature (Aabø, 2009;
Imholz & Arns, 2007; Matthews, 2011) to explore the field of library valuation; that is, to
obtain an estimate of the monetary worth of libraries, framed as a “return on investment”
(ROI) for taxpayers. Aabø’s (2009) preliminary synthesis of valuation studies returned
the following conclusion “for each dollar of taxpayers’ money invested in public libraries,
the libraries – on average – return a value to the citizens of four to five times more” (p.
322), which is consistent with the result of $4 to $6 of ROI presented in Matthews
(2011). While both Aabø (2009) and Imholz & Arns (2007) caution that the field of public
library valuation was “young” at the time of their reviews, publications in peer reviewed
journals since these reviews were published are limited. Additionally, a number of these
more recent articles are focussed on mitigating identified shortcomings of a strict cost-
benefit analysis. McIntosh (2013) attempts to address the wide variation in estimation of
cost savings to the patron used in the cost-benefit calculation by using a contingent
valuation approach. De Leon (2021) evaluates a Social Return on Investment (SROI)
EXAMINING THE “CULTURE OF ASSESSMENT”

metric in a public library setting, with the aim of quantitatively accounting for “social
values created by activities and investments” (p. 177). However, Matthews (2011)
indicates that the utility of such valuation studies is limited when applied to justification
of library budgets; the ROI figure may not be sufficiently “compelling” (p. 10) on its own.
As a cumulative representation of library services, it also does not provide sufficient
information for libraries themselves to make decisions about their operations; none of
the review articles (Aabø, 2009; Imholz & Arns, 2007; Matthews, 2011) provided a
synthesized breakdown of the valuation for individual programs and services. In the
corporate context from which it was drawn, ROI is appealing as a simplistic cost-benefit
topic to compare the potential or realized profitability of different projects while
abstracting away from the details of those projects (Phillips & Phillips, 2006). The
significant work done in the field of library valuation and library ROI as
described here indicates a similar desire for a simple measure of “value”. However,
even in the corporate context proper use and interpretation of ROI can be nuanced
(Friedlob & Plewa, 1996) and it is not self-evident that a metric designed to compare
profitability between revenue-generating enterprises is appropriate for use here.

A Broader View: Outcome Assessment

Indeed, as this limited applicability of strictly economic measures was recognized
early in the literature (Usherwood, 2002), significant attention in the past two decades
has also been given to providing libraries with methodologies to support assessments of
library services and program outcomes. Outcome assessment has multiple different
definitions within the literature (Lyons, 2012), but they are consistent in their focus on
benefits to the library patron (e.g., parents who bring their children to story time see
increases in early literacy) versus benefits to the library (e.g., the library sees an
increase in circulation of children’s books). In particular, the Public Library Association
(PLA), a division of the American Library Association (ALA), has created a series of
resources to assist public libraries with defining and executing outcome assessment
programs (American Library Association, 2021; Nelson, 2008; Rubin, 2006). These
resources, among others (Lyons, 2012) advocate for a basic cyclical assessment
process which begins a planning phase that identifies overarching values and
principles, as well as goals and objectives. The intention is that these planning steps are
to be performed with input from the community (Nelson, 2008, pp. 35-42; Rubin, 2006, p.5). From there, a gradual shift in responsibility to library staff occurs to define the possible service responses and/or programs, the metrics to collect and the subsequent data analysis. Identified metrics for collection and analysis can take a number of forms, including:

- Direct measures of the program or service being evaluated, such as number of attendees, program cost or participant questionnaires (Rubin, 2006).
- Proxy measures, including library-oriented statistics such as regular expenditures, number of staff, collection and circulation counts, patron visits, reference transactions and numbers of programs offered. In the U.S., these statistics are also collected nationwide by the Institute of Museum and Library Services (IMLS) and reported in the PLA’s benchmarking report (Public Library Association, 2021).

While it is recognized that the “benchmarking” statistics listed above are insufficient for adequately evaluating program or service-level outcomes, direct measures that are quantitative in nature are often prioritized over qualitative approaches in professional practice publications. Rubin (2006) defines program indicators as “a statistic that summarizes the user’s achievement of an outcome” (p. 33). Indicators in this framework are necessarily paired with a target, defined as “the proportion or quantity of participants that must meet an indicator for the program as a whole to be considered successful” (Rubin, 2006, p. 38). This strategy is consistent with the corporate quality management framework discussed in the Introduction.

However, it was difficult to find purely quantitative studies in the recent literature using this or similar approaches. Current published quantitative efforts (Chow & Tian, 2019; Crawford & Syme, 2018; Mathiasson & Jochumsen, 2020) seem focussed on reviews of existing library records as described in Rubin (2006, pp. 45-46), with an emphasis on novel data analytics techniques drawn from the “big data” field. For example, Chow & Tian (2019) investigate predictive relationships between circulation statistics and community “quality of life” indicators such as educational level, median household income, and number of jobs.
Concerns with Strictly Quantitative Approaches
Numerous articles have highlighted difficulties with a strict preference for primarily quantitative studies (Closter, 2015; Irwin & Silk, 2019; Irwin & St-Pierre, 2014; Lyons, 2012), which may explain the recent shift away from such approaches in the research. The leading concerns center around (a) an awareness that the quantitative approach can only measure what is measurable, and may be missing crucial context (Closter, 2015); and (b) the quantitative results may lead to inappropriate generalization of the conclusions (Lyons, 2012). Specifically, evaluations designed to use quantitative outcomes to represent overall library value without considering social impacts can result in libraries “exist[ing] in a correlation trap of their own making” (Irwin & Silk, 2019, p. 325) when stakeholders directly correlate metrics associated with number of card-holders, circulation statistics and program attendance (simply because they are measurable) with the quality of public library service outside of the context in which those metrics are produced, changes in those metrics may lead to incorrect conclusions about the value of the library to the public. Rubin (2006) explicitly points out that the method proposed in her book is not intended to use the “systematic methodology or standardized framework of social science” (e.g., there is no formal sampling procedure, or control group) and therefore it is not possible to discern if change in the specified indicator is as a result of the program or service, limiting the use of any quantitative predictions from the results in future planning. As such, many recent studies on outcome assessment in public libraries make use of either mixed-methods or primarily qualitative approaches, a summary of which is given in the next section.

Current approaches to library assessment
Given the potential gaps introduced by the focus on primarily quantitative studies, a frequent remedy is to adopt a mixed-methods approach, where both quantitative and qualitative tools are used for evaluation (Armas, 2019; Becker et al., 2009; Graham & Gagnon, 2013; Greene & McMenemy, 2012; Nwofor & Chinyere, 2015; Oliphant, 2014; Pabērza, 2010). The most commonly used tool is a questionnaire with both closed- and open-ended responses to allow for both the calculation of statistics and deeper dives into the context. For example, Pabērza (2010) used this approach (augmented by focus groups) to assess public library usage and value perception in Latvian public libraries.
Oliphant (2014) performed a comparative study on impressions of public library value in Alberta from both users and non-users via phone surveys that asked both quantitative questions such as how much the participant would pay for an annual membership fee (p. 353) and qualitative interview-style questions such as the participant’s perception of library staff (p. 357). Armas (2019) also used a combined questionnaire and in-depth interview strategy to demonstrate the impact of the public library on the lives of older adults in Cuba.

Literature on mixed-methods approaches of the type described above tends to focus on larger populations, due to the need for sufficient data collection for statistical analysis. There is a recognition that scenarios involving smaller groups of participants (such as local programming) may be better served through primarily qualitative study designs such as case studies (Gahagan & Calvert, 2019; Streatfield et al., 2019; Teasdale, 2020), focus groups (Bishop et al., 2003; Elbeshausen, 2007), expert panels (Halpin et al., 2015) and in-depth one-on-one interviews (Aptekar, 2019; Griffis & Johnson, 2014; Johnson, 2012; Liew et al., 2021; Valdivia & Subramaniam, 2014; Vårheim, 2014). As an example, Liew et al. (2021) use interviews with a small population (where “small” in this instance refers to a sample size of 8) to conduct a phenomenological study on digitized Indigenous knowledge collections that spotlight key cultural values that “have not been embodied in previous value-impact frameworks” (p. 1588). Public library makerspaces are identified as an area where the heterogeneity of maker usage patterns makes a quantitative assessment of outcomes difficult, leading researchers to pursue study designs that include participant-defined success criteria (Teasdale, 2020) and “storytelling” by librarians observing the makerspace users (Gahagan & Calvert, 2019). Evaluation of the social capital created by libraries is a clear topic of interest in the literature, with both patron and library staff interviews forming the core of the study methodology (Aptekar, 2019; Griffis & Johnson, 2014; Johnson, 2012; Vårheim, 2014). Action research (MacDonald, 2012; Streatfield et al., 2019) also provides a potential avenue for exploration; an example of this approach is given in Bishop et al. (2003) where a participatory action research (PAR) framework was used to build a collection of digital tools and resources for women’s health in collaboration with
African American women and other community partners. Through methods such as the use of discussion groups to build scenarios and a community action plan, and situated user evaluations, local knowledge was placed at the center of the design process and assessment was presented as a collaborative concept useful throughout the project lifecycle that built capacity within the diverse community as they worked toward social change. Streatfield et al. (2019) similarly used PAR as a key methodology in the Global Libraries Initiative.

It should be noted that these qualitative approaches are identified as possible tools within practice-oriented handbooks like Rubin’s (2006), despite the upfront focus on quantitative indicators and targets. ISO 16439, the standard for impact assessment in libraries (Creaser, 2018; International Standards Organisation, 2014), indicates that the recommended methodologies (both quantitative and qualitative) are “most effective in demonstrating impact when used in combination” (Creaser, 2018, p. 90). As mentioned previously, the difficulty is that, regardless of context, quantitative measures are often presented as superior to qualitative ones, despite the noted shortcomings and the potential loss of opportunity to collect a richer data set (Gahagan & Calvert, 2019). Studies making use of primarily qualitative approaches often contain justifications for why a quantitative study was not performed, indicating a belief that quantitative approaches are the default, and deviations must be accounted for.

In summary, it is clear from the literature that a need exists for public libraries to have some kind of evaluation system in place. Even if it was possible in the present political climate to successfully argue that libraries represent a public good and are therefore entitled to a portion of public funds independent of any evaluation processes, decisions would still need to be made about how to allocate those funds: staffing, collections, programming, infrastructure (and its maintenance) among other potential sources of expenditure. An appropriate, effective evaluation system can help library staff and administration make decisions that are congruent with the needs of the community and the library’s stated goals (Magnus et al., 2018; Nelson, 2008). This library and community-centric use of the tools of assessment can be clouded, in practice, by the neoliberal narrative wherein the public library is a “quasi-market”, and the patrons are “customers” (McMenemy, 2009, p. 400). As has been previously
discussed, quantitative methodologies which can provide “performance indicators” can be elevated within this narrative above qualitative approaches that may indeed yield the information necessary to make appropriate community focussed decisions.

Therefore, the underlying premise upon which the “culture of assessment” has been promoted has introduced two paradigms that must be addressed: (a) despite widespread use (and endorsement in professional practice publications) of a variety of assessment tools, the available options are placed into a “validity” hierarchy with quantitative tools at the top level, as discussed in this section; and (b) the need for outcome assessment is typically presented as a threat-response to the very existence of public libraries (Farkas, 2013; Greene & McMenemy, 2012; Lakos & Phipps, 2004). This second point will be discussed in detail in the next section.

The Future of Outcome Assessment in Public Libraries

Lakos & Phipps (2004) state that “to successfully respond to threats to their survival and demands for better, faster delivery of information, libraries have to incorporate assessment into their everyday activities” (p. 351). This kind of framing is in stark contrast to the activities suggested by the literature as necessary for successful implementations of a “culture of assessment”. Common guidance includes establishing a formal assessment program where the requirements and goals are clear to all involved, involving external experts, providing training in research methodology and assessment techniques, implementing a decision support system, and defining a pathway whereby assessment outcomes are available for use in strategic planning and programming (Farkas, 2013; Hiller et al., 2008; Irwin & St-Pierre, 2014; Lakos & Phipps, 2004; Lyons, 2012; Rubin, 2006). The common thread in all these recommendations is that they all speak to lengthy processes that will require both time and funding; it is not surprising that presenting them within the anxiety-producing context of impending threat to survival has not produced the desired engagement. Farkas (2013) calls this to attention specifically in the discussion of needing to create “a sense of urgency” when promoting organizational change, “Urgency is not driven by fear or anxiety, but by opportunities and a sense of possibility. Therefore, a change message designed to create a sense of urgency should not be fear-producing” (p. 19). Interestingly, the
solution to the lack of engagement with assessment programs is presented as a need for cultural shift within the population under threat, complete with statements from management about how the staff are unwilling to engage in additional work (Irwin & St-Pierre, 2014, p. 7). It is not clear from the literature if this *a priori* assumption that culture change of the type described is required when building assessment programs was ever challenged, or alternative means proposed.

Although many provide commentary, some studies (Düren et al., 2017, Greene & McMenemy, 2012) have engaged directly with the question “What is the impact of the neoliberal ideology on public libraries?” Greene & McMenemy (2012) demonstrate via a mixed-methods discourse analysis that the introduction of neoliberal language reinforces “negative stereotypes of a service in distress”, which can only be addressed via ensuring “value for money” (p. 34). Düren et al. (2017), despite a positive concluding evaluation of NPM, were not able to demonstrate that the adoption of NPM ideologies (and corresponding assessment approaches) were successful in preventing the closure of libraries. While the number of these studies directly addressing the outcomes of the introduction of NPM into public library management are limited, they do indicate a need to explore the applicability of the underlying premise upon which NPM is based to the public library context. It is not apparent from the present review of the literature, even key foundational papers advocating for the “culture of assessment” (Lakos & Phipps, 2004), if the question “Is a corporate quality approach to assessment appropriate for a non-revenue-focused public organization?” had been asked, or the conclusion simply taken as axiomatic.

**Conclusion**

In conclusion, there is ample evidence in the literature that stakeholders and decision makers agree on how the public library provides value, and that librarians and LIS researchers can use a wide variety of study methodologies to establish that value at multiple population scales from individual groups at local branches to multinational efforts. What is less apparent is if these efforts are having the effect promised by the culture of assessment narrative. Are public libraries in general seeing relief from the impending threats of closure as a result of determining their ROI, measuring their benchmarking statistics and creating balanced scorecards? Are public libraries better off
for having started to think of their patron community as “customers” and evaluating programs and services in that context? These are open research questions, as is the following: is it time to reclaim the assessment narrative from the capitalist context in which it has been placed, and put it back in service to public librarians themselves, so that they may in turn serve their communities?

References


https://doi.org/10.1353/ils.2013.0013

https://doi.org/10.1108/s1876-0562(2012)0000006005

https://doi.org/10.1177/0961000612470277

https://doi.org/10.1177/0961000613497746

https://doi.org/10.1108/14678040810928444

https://doi.org/10.1300/j118v26n03_03


https://doi.org/10.1080/01616846.2019.1571398

https://doi.org/10.1177/2158244014561214


EXAMINING THE “CULTURE OF ASSESSMENT”


