



Discussion Paper

Towards a Statistical Index for Ontario's Public Libraries

Jan. 22, 2016

Authors:

Stephen Abram, MLS

Robert Molyneux, MLIS, PhD

ONE VOICE FOR ONTARIO PUBLIC LIBRARIES



Discussion Paper Introduction

Stephen Abram, MLS, Executive Director

This editorial introduction is an overview of our process at FOPL and call for volunteers to contribute to the development of a province-wide approach to data which will make our systems increasingly successful.

The Federation of Ontario Public Libraries has been investing over the last 24 months in statistics for Ontario's public library systems. To this end we have:

- Participated in CLA task force on national statistics
- Hosted 3 iSchool symposia on future measurements for libraries
- Lobbied for open data for Ontario public libraries (win!)
- Published a longitudinal analysis of Ministry data collection for 2001-2013.
- Published Market Probe opinion polls for 2015 (building on 2001, 2006, 2010 polls)
- Hosted and recorded several webinars about Statistics and Measurements in public libraries.

In 2016, our goals are:

1. To continue the discussion of relevance and timeliness of Ontario public library comparative statistics.
2. Publish an updated longitudinal analysis of the data collected from us by the Ministry for 2001-2014 in early 2016.
3. Start the development of an index to compare libraries in Ontario on 21st Century strategic benchmarks that align with the role of libraries beyond just circulation and gate-count.

We are very pleased to be working with library stats expert, Robert Molyneux, MSLS, PhD, for his excellent analytical skills and perspectives in developing our reports and his contributions to this discussion paper.

It is our hope that this paper and our statistical analyses will serve as an important launching point for broader discussion in our community on what WE NEED with respect to statistical data, provincial benchmark measurements, and historical context for our annual strategic planning reviews and new strategic plan development.



Combined with the completion of our 2015 Market Probe Canada Public Opinion Poll on Ontarians attitudes and Usage of Ontario's Public Libraries, we have updated the FOPL public opinion polls done in 2000, 2006, and 2010. Each of these polls has provided insights into our operations and how we remain successful at fulfilling our public mandates.

I am excited that our analyses show that public libraries are at two tipping points. Traditional library measurements peaked or plateaued in 2010. Modern library measurements that include our emphasis on programs and digital usage show that we are on a new arc and digital usage appears to be exceeding in-person usage for the first time in 2015. It is truly an exciting time for our important and value-laden sector.

FOPL has been very active these past two years. The membership investments in research, marketing and advocacy on a collaborative basis are bearing fruit at costs that would be unaffordable to individual library systems. The return on investment for your membership in FOPL has never been stronger! We can report significant progress in all three Federation of Ontario Public Libraries strategic objectives in 2014: Research and Development, Marketing, and Advocacy.

This discussion paper educates you in the background on indices as they apply to libraries and how they can be used for analysis, ranking and comparisons of Ontario public library performance from 2001-2014. Credit is due to our Ontario Ministry of Culture, Tourism and Sport for jumping on the open-data trend and making the raw data that public libraries have contributed since 1998 available for further analysis and usage, and we thank them for their efforts. The Ministry has provided some funding in 2015 in addition to FOPL funds to this project to update the data, and to make some comparisons to the public libraries in US states and Canadian provinces. This is great and our sector's data is now fairly current for the first time in many years.

I believe that that these projects project provides strong value for public libraries from the data our sector has provided for many years and underpins communication of our value that show the Ministry and the public the strong impact of public libraries from their long-term investment in our libraries.

What you need to do:

Simple really, we need to discuss this discussion paper. We need a dedicated cohort of library leaders to understand the issues and make progress. We will have future meetings, webinars and teleconferences on the work started here. Please VOLUNTEER!



2025 Summit). To this end we will have much of the research and information we need to plan for future needs. We will have:

- Longitudinal Statistics Data on Ontario public library performance (2001-2014)
- 2015 Data on our marketing channels, market positioning, and preparedness for a 'push'
- Insights from stakeholders in libraries, boards and municipalities.
- A 2015 public opinion poll of Ontarians attitudes towards public libraries and how they've changed from our 2000, 2006, and 2010 polls.

We are better prepared than ever to do local system strategic planning as well as engage in province-wide collaborative strategies. It's an exciting effort and we're getting there. We're trying to be strategic, plan well and provide you with the information you need to plan locally in your community. We're here to support you.

Stephen Abram, executive director

FOPL: 416-395-0746

sabram@fopl.ca

ONE VOICE FOR ONTARIO PUBLIC LIBRARIES

Towards a FOPL Index

This document is the latest in the series of FOPL investigations using the Ontario public library data compiled by the Ministry of Tourism, Culture and Sport. The purposes behind this research is to use these data to help librarians in practical ways.

Previous reports analyzed individual variables on different sets of the Ontario libraries and for two different sets of years. The goal was to get a general picture of these libraries based on the data and to examine the data themselves. A conclusion reached about the data was that another year or two of collection and reporting should result in a more mature dataset with fewer of the seemingly improbably high or low values which we found here and there with the current data. Given that the data are now available in a format which allows ready manipulation, these anomalous odd values will be noticed and examined. Data improve with scrutiny and use and as odd values are questioned, reporting of data gets better. A result from this kind of process will be a set of data which can be better used to assess strengths and weaknesses of these libraries. Such a dataset would also allow for benchmarking the state of Ontario's public libraries in a year or over time.

Overview

We have continued working on these data in hopes of developing more useful tools to examine the Ministry's data. The Ministry's library datasets are rich but certainly complex. Finding out the important stories lurking in these data requires work.

We present here the results of an experiment constructing a means of assessing the strengths and weaknesses of the libraries using the latest available data—those for 2013—and attempt to provide a useful index of the results. We eliminated those libraries which contract for services with other libraries because their data are not separately available in detail but are included with the data from the libraries they contract with. We have, then, 311 libraries reporting in 2013 and those are the data we use.

In the second section below, we have a brief review of the history of such indexing efforts but one, the BIX, gives us a working definition that guided through the experiments outlined here:

The goal of BIX is to effectively describe the performance of libraries with statistical data, and thus to assist evidence-based communication between libraries, their funders and policymakers through transparency of services and to enable libraries to assess their strengths and weaknesses and indicate opportunities for quality improvements.

<http://www.bix-bibliotheksindex.de/en/project-info/general-information.htm>

We will review previous assessment and indexing efforts below but a common characteristic is that they combine an inventory of quantitative measures from the libraries and then use these measures to assess the performance of libraries.

It is hoped that a critical discussion about what we have done will lead to an index which can be used by FOPL members to assess their libraries, improve their services, and provide evidence useful for in budget discussions.

Organization of this paper

In the first section, we discuss the spreadsheet (Table 1) which presents, as its title indicates, the results from our approach to an Experimental FOPL Index. What we have done could also be called an assessment inventory by ranking 16 traits of Ontario's libraries. In a broader sense, it is a template for such an index because we can easily substitute other variables which the members might prefer. History and local practice will constrain such decisions.

For this initial—and non-public—distribution, we include Table 1R which differs from Table 1 in that it contains a summary ranking of all libraries by their size groups using the same data in Table 1. We discuss Table 1 first and then will discuss Table 1R and the differences between the two.

It is quite possibly premature to make Table 1R widely available for reasons discussed below. This question should be considered: publish Table 1R or not?

The second section is larger as it delves into methodology used to construct this index and necessarily includes a review of other library data indexes which informed our work. Table 2 maps the data in this Experimental FOPL Index to these earlier indexes. The reader will note that a number of the variables appear in more than one of these indexes. In fact, we used many in our earlier work because they are commonly recognized as measuring important aspects of libraries. We also discuss the fact that a review of other public library data collections in other countries and Canadian provinces shows many possible ways to measure libraries.

The term “index” has occasionally been a contentious one in the library field. There are “indexes” and there are “indexes.” This is a most interesting

question but not central to our purpose here but worth a bit of a look in this second section.

Developing such an index is an attempt to measure the strength and weaknesses of libraries using quantitative measures: data. Quantitative measures have their own strengths and weaknesses. One of the most used quantitative measures is circulations per capita which often serves as a proxy for a library's activity. If we analyze a set of libraries, generally we will agree that it is better to have more circulations per capita than fewer.

In looking at this preliminary work, there will be readers who see libraries showing better or worse results than expected and often because of qualitative criteria. Standing in a library and looking around is informative but hard to measure in a comparable way across all the libraries in Ontario. Quantitative measures are what we have and we start by using them to assess libraries.

Quantitative assessments are important but in thoughtful moments, away from the day-to-day, one is likely to acknowledge that qualitative assessments may give us a better picture of the actual state of a library. Quantitatively, we have counts of the number of programs offered by libraries and can even calculate the number of programs offered per capita—per resident population or even, now, per registered borrower. However, not all programs are the same. A story telling program presented by a great story teller will have more effect on young (potential!) readers than one not so skilled. A knowledgeable observer will see the differences between effectiveness of a program offered by story tellers with different skills. However, we do not measure and report these kinds of qualitative aspects of libraries in any comparable way. As a result, we have rigid, seemingly precise numbers to rank and assess libraries as we present here but always the conclusions we draw have to be tempered with context and judgment. And that is so particularly because of the anomalies in our data.

The third section is a deeper probe into the methodology used here. This section is rather detailed. Table 3 has formulas used to operationalize the variables. It is one thing to say: circulations per capita but which of the Ministry's data elements were used to create the values used in Tables 1 and 2? For those happy few who hear the beguiling siren call of library data, and who might want to help, this section will be important. The work presented here is the result of many complex operations using spreadsheets and help in checking these calculations would be appreciated.

SECTION 1: AN EXPERIMENTAL FOPL INDEX/ASSESSMENT INVENTORY

Rather than start with the methodological discussion first—as fascinating as it is—we present the results of our work in a spreadsheet, and invite the committee to try it out. Given that it is a live spreadsheet, remember that spreadsheets are tricky and caution is in order: you can make an unseen mistake in seconds. Save a copy a safe place and then work off a copy of that safe version. Working with spreadsheets is given to inadvertence so you will have something to start over with when—not if—you need to. Fair warning: when.

The spreadsheets in this paper are in Excel (.xlsx) format. Libraries which do not wish to purchase Excel will find that LibreOffice suite includes a spreadsheet program called Calc which is an excellent alternative. (<http://www.libreoffice.org/>) LibreOffice is an open source project and the software is free.

Table 1 summarizes the results from other Excel tables which have the underlying detailed data and their manipulation. Those underlying spreadsheets will be discussed in the Section 3 and will be available to members who wish to examine them. Table 1 presents our approach to an Experimental FOPL Index. It is also suggested as a robust design given that it includes most of the important aspects of other such library indexes while being easily adapted to the member's requirements.

As is common with library data and for reasons discussed in the Primers, there are five “cohorts” of like-sized libraries. The skewed distribution of libraries by size is a factor we have to deal with if we are to understand our world and separating libraries by size for analysis is a recognized method. The largest libraries exist in a different environment from our smallest libraries. We can acknowledge both those differences and, at the same time, acknowledge our similarities and all work towards common goals.

Continuing long practice, we use per capita (the ratio of the variable in question divided by the resident population) and other similar ratios. The five cohorts are classified by “resident population.” The first cohort includes the 27 libraries with a resident population greater than 100,000. The four others which each have 71 libraries and include the First Nations libraries—unlike our previous studies. We do not separate the First Nation's libraries into a separate group here as we have done previously. These cohorts are arbitrary but mathematically convenient. Cohorts can be designed

based on other criteria than size such as local libraries, consortia, region...to name just a few.

The spreadsheets report the ranks of these variables for each library within its cohort. There are 16 variables grouped in “dimensions.” These dimensions derive from the BIX (discussed below,) which was developed for German libraries, and reflect a fact that libraries do many things which can be grouped by function. The next section will discuss these dimensions more but the reader will see they reflect aspects of working libraries. However, the method creating these rankings within the cohorts makes accurate comparisons between libraries in different cohorts using the data in Table 1 almost impossible. It is best to compare libraries within a cohort or redo the cohorts to group libraries which are alike by other criteria.

Table 1 orders the libraries in alphabetically. The thinking behind this order is that in this experiment, librarians will want to look for their libraries or other libraries they are familiar with. The members can sort them differently if they choose.

This spreadsheet is two things: a serious attempt to create a useful method for Ontario’s librarians to assess their libraries through varying “dimensions,” and a quest for help to make such a thing better.

Let us take a tour of Table 1. We have the promised 16 variables as noted grouped within the four dimensions borrowed from the BIX. The four dimensions are:

Service

Usage

Efficiency

Development

And each has a set of variables which, it is hoped, fit within the dimensions.

The variables we use are largely the BIX variables but not all. We have three reasons not to do a direct emulation, were we so inclined. One is that the Ministry doesn’t always collect the same data as those used there. Second, the BIX English pages are helpful but not unambiguous so we could not always tell if we had the same variable. The third reason is that we have also tweaked these data to make it possible to compare the results to the LJ Index (discussed below) should anyone be interested. There is more discussion of this matter in the next two sections. However, a great strength of the BIX is that it has been developed and tested over time.

Starting from the left, we have the Ministry's library number, then the name of the library and its library service type. Next is our old friend, resident population as reported this year, 2013.

We calculated the 16 variables for each of the libraries within their cohorts in the base spreadsheets. There were one or two variables per spreadsheet with the libraries ranked by the variables. The "best" level of the variable (highest in most of them but in a few, the lowest) was ranked 1, the second library, 2, and so on.

Then we created four summary spreadsheets of each of the dimensions and created an average of the ranks for each library, within each dimension. Ranking the variables within each cohort and dimension in order and calculating the average rank gives us a rough draft of rank for that library in each dimension. Our assessment inventory of library performance was growing. These dimension tables have the libraries in alphabetical order within the cohorts.

For Table 1, we move beyond those foundational spreadsheets and move closer to our goal. We summarize the results of the spreadsheets for the four dimension in Table 1 by calculating the arithmetic mean of the 16 ranks.

There are three variables in our spreadsheet with asterisks. These are variables in which the ratio being ranked is inverted, that is, the better rank is has a lower value. With circulations per capita, a higher number is better than a small one. If you seek to compare the number of circulations per dollar of acquisitions expenditure, you probably want the lower number. If there is \$1 spent on acquisitions expenditures is related to 10 circulations that is better than \$10 spent for 10 circulations. It is regarded as more efficient.

Column G is the average of the 16 ranks for each library across all four dimensions. Arguably, this kind of broad, systematic analysis, will give us a broad, systematic picture of the state of Ontario's libraries.

What we have with the average ranking is just the unweighted average of all 16 ranks. Weighting is a technique we discuss in Section 2. It is a way to recognize that some variables are more important than others and FOPL members may wish to weight some variables.

If we totaled the ranks, each library would have the same order with averages as totals given that they are mathematically equivalent.

The next two columns should also be useful. We have the average rank for each library by this method but what are the highest and lowest values for each of these libraries across these 16 variables?

Note the Highest Rank (Grey) and Lowest Rank (Green) which give us the range of these ranks for the libraries. In order to see how one might use the highest and lowest columns, let us take a look at two libraries in the first cohort.

London Public Library has an average rank of 8.2 while Essex County's average is 20.0.

London's highest rank (see the gray column) is 1 and examining the ranks we see it led the other libraries in this cohort in Program Attendance per capita, Resident population per workstation and Visits in person per capita. The lowest rank for London PL is 22nd and that is "Collection units per capita." It's collection size per capita, relative to the other libraries in this cohort, is relatively small.

What about Essex County? What can we glean from this spreadsheet? Curiously, Essex County's rank for "Collection units per capita" is highest in this variable for the 27 libraries in this cohort! In Essex County, then, there are many more "units" per capita than in the London Public Library. Note that the definitions of these variables are in Table 3 where a "unit" as analyzed here is defined.

Is this particular number 1 ranking good or bad? Well, there is a reason we talk about "weighing" (as distinguished from "weighting" variables by importance) evidence where we have a situation where one might say "on the one hand...but on the other."

Without standing in the libraries and getting to know them, we cannot say what is going on definitively but we can examine what we have here and speculate about the next set of questions to ask based on the reported data.

Even though Essex County has the largest collection per capita, note that the circulations per capita and the stock turnover (average circulation for each item in the collection) are ranked last in this cohort. Essex County has a relatively large collection but relatively low circulations—compared to the other members of the cohort. London with its lower ranked collection units has a higher ranked circulations per capita (10th) than Essex County and an even better ranked stock turnover (7th.)

As we warned: this result is a speculation and one would always have to go back to the raw data...these results are an index to those data but the raw data are closer to an answer and visiting the libraries would be better. However, the better the data and the better we design such indexes, the better we will be able to assess any group of libraries. Then we can better probe the imbalances Hennen discussed which we review below.

Similar kinds of stories are there for examination. What is our library's highest rank? Lowest rank? Do these results seem right? How can we improve the results? How do we compare to other similar libraries?

Table 1R

Table 1R is included in this preliminary distribution and it differs from Table 1 in two ways. First, there is an added column highlighting each library's rank using the average of all ranks. We also include the sum of all ranks. The average is this number divided by 16. This column is just another view of the same data. Second, the libraries are in order by this rank. Ties are resolved, that is, if two libraries have the same score, they are the same rank in the table and the next library will be where it would have been ranked without the ties. In Cohort 1, Cambridge, Oshawa, and Whitby are tied in 11th place. The next library, Greater Sudbury, is ranked 14th.

Readers of those tables will note some of these "rankings" may contradict their informed knowledge of various libraries. These contradictions will largely be a result of those anomalous values mentioned above.

How did these anomalous values arise? There are many origins but one can imagine a librarian pressed for time trying to fill out the detailed forms putting down a quick answer, rather than an accurate one. As we have pointed out, it takes several years to discover these kinds of anomalies out and improve them for the next publication.

The contradictions present a dilemma for consideration. On the one hand, publishing the summary ranks will be embarrassing to some FOPL members. In the data world, it is a common notion that data collected for one purpose may be reused for another but these data have not been used for the purpose at hand and there could be unwelcome surprises.

On the other hand, by publishing Table 1R, the data problems will be unmasked, and, in time, the data will improve. Given that the 2014 data are on the way, such improvements will not be seen until the 2015 data--at the earliest.

SECTION 2: BACKGROUND OF RATING LIBRARIES USING DATA

Indexes are constructed in different ways and to different ends so the term “index” might almost be regarded as a linguistic flag of convenience. We have offered a more accurate alternative of “assessment inventory” to describe the underlying processes and results. The BIX goal we highlighted earlier can be used to describe a variety of indexes from stock markets, currency, or libraries, with appropriate changes in what is being measured. There are methodologically narrowly defined indexes like the obsolete ARL Library Index (discussed below) and broader ones like the BIX.

Given the audiences they are designed for and the characteristics of the data, the construction of indexes will vary. In the case of the BIX, the HAPLR, the LJ Index, or a FOPL Index, are assessment inventories. That is, a set of library traits and with libraries ranked by these traits to the end of providing a tool for assessing libraries.

Historically, our colleagues were as interested in comparing their libraries as we are. A library history course will likely discuss the rivalry between the libraries of Alexandria and Pergamom with fragmentary reports of their reporting on how many scrolls they held and, naturally: more were better!

More recently, the 1835 report by Antonio Balbi was the first careful investigation of the rating question. Balbi assembled data on the size of libraries in the largest European cities with the hope he would be able to prove that his adopted home city of Vienna had the biggest and, of course, the best. He used unsystematic accounts of volumes reportedly held by those libraries from a variety of sources which had been published over a number of years. Balbi, sadly, was “disheartened by the disparity of opinion” he had from these remarkably varying published reports of the size of the European libraries. He went on, as is sometimes said, to make lemonade out of those disheartening lemons. His seminal analysis of the problems with library data systematically lays out the problems we face even today. There has been limited advancement in library data analysis since his insights.

However, we have great advantages over Balbi: with the Ministry’s work, we have systematic, annual data collection which is professionally compiled and which reports on many more aspects of modern libraries than volumes held. That advantage gives us an opportunity to create measures of libraries with more—and newer—aspects of libraries than just volume counts. Also, we can use these measures over time.

Measuring libraries by a consistent set of variables over time allows us to examine trends and find out what changes help libraries improve services for their users.

There are five indexes we considered as models in our experiments:

The ARL Library Index

The Association of Research Libraries (ARL) data continue a series started in 1907/08 by James Thayer Gerould. As a result of that early start, the ARL series appears to be the oldest continually collected set of library data in existence. The early years of the data showed a clear interest in the size of these libraries in volumes held and this interest continued for years.

These research libraries are major institutions comprising academic, government, and city libraries in Canada and the US. The question of membership criteria arose over the years as the Association grew larger. How was membership to be decided? Which library was a research library and eligible to be a member? What is a “research library?” Kendon Stubbs, probably is the best Statistical thinker the library world has produced, developed this Index to quantify the criteria for membership based on a Factor Analysis of the ARL member data. It was the most sophisticated approach and probably the “right” one for any such index from a methodological standpoint. It has much to commend itself but it is one hard for our community to understand. The Association no longer uses it.

Factor Analysis is a mathematical method to examine the many variables ARL collects and to pick out the major aspects of these libraries in the data. The method will provide a sparse index. That is, if there are several variables which are highly correlated, why include all of them because one will be enough to describe that aspect of the libraries?

For instance, in the 2013 Ontario data, the correlation between the Total Operating Expenditures and Resident Population Served is 0.98 with 1.00 being the highest positive correlation—that is, the two are almost the same for many analytic purposes. A number of library variables are similarly highly correlated and in a Factor Analysis, we can boil these correlated variables—set by some threshold—down to one variable, we have improved our model. After this kind of analysis, we will have mathematically identified the few “vectors” in these data.

For the ARL membership question, Stubbs's solution is elegant. However, for a library audience the results are hard to relate to anything on our world. If one is interested in assessment with an idea of trying to see what works best in managing libraries, collapsing variables in a sparse way, throws away potentially useful information. As we will see, this is a point Tom Hennen made in compiling the HAPLR which, like our Index, was not sparse.

We will put this method aside for now. It may well be of use in discussing matters of information policy for Ontario but not for this stage.

US State Ranking Tables (USSRT)

US agencies have published compilations and analysis of public library data. The most popular report is a series of (now) 22 state rankings by various variables most of which are per capita. These are based on summary data for the 50 states plus the District of Columbia which is treated as a state for purposes of this analysis.

These 22 variables have been tracked at the state level for a number of years. Little work has been done in analyzing trends in these data or analyzing states by these ranks. Even so, this compilation is immensely popular with the state librarians. In the US, each of the 50 states has a "State Library." State libraries are required for certain federal programs such as grant funding. Their scope, task, and names of these "libraries" vary considerably outside the narrow federal requirements.

In earlier FOPL reports, we attempted to emulate the rank order tables on several occasions with all the Ontario libraries and it seemed that direct emulation of this method developed for 51 states did not seem to scale to 300+ Ontario libraries. It is one thing when the ranks are 1 to 51 versus 1 to 311.

However, still open is the notion of summing ranks—that is, the analysis of multiple ranks. Suppose, one library is ranked first on 5 categories and another last, the sums of those ranks would indicate one was ranked higher than the other. If the criteria used for that ranking were reasonable and the variables related to what makes a library "better," then we would have a case that the better ranked library was, in fact, "better." And that leaves open the question of what variable tell us what is "better."

In fact, this method is pretty much what library indexes do.

HAPLR (Hennen's American Public Library Ratings)

<http://www.haplr-index.com/>

Tom Hennen was a working librarian who started the HAPLR to help him and others run their libraries and he published the ratings from 1999 through 2010. His interests were practical as the reader can see by looking at the HAPLR Website. Although it appears the HAPLR is moribund, the site still presents the output of a productive and thoughtful mind interested in practical library problems and their measurement. It is a rich resource which repays study.

HAPLR used 16 variables including both inputs (resources made available like revenues from various sources) and outputs (what happened as a result of those resources.)

HAPLR was attacked rather sharply, particularly by Keith Curry Lance, a well-known and respected library data analyst and later by Ray Lyons, another respected analyst. Both also have years of experience working in library data.

Eventually, they produced the Library Journal LJ Index of Public Library Service which we discuss below but we will get a bit ahead of that story now. There were several points of their attack and rather than reprise the entire and interesting back and forth between the three of them, there are three points of criticism that concern us now.

The first criticism: the fact that Hennen used correlated variables among his 16 variables. In our broad use of "index" here, given his purpose of providing a practical tool for working librarians, his "Ranking" did not exclude correlated variables.

Second is the question of weighting and the subjective aspect of it. HAPLR weighted variables. As we mentioned, weighting is a technique which, in effect, recognizes that some things measured are not as important as others and mathematically gives varying importance.

How does one do this kind of weighting? Largely, it is subjective in our world. In the case of HAPLR, Hennen polled librarians on a public library listserv and used what he gleaned in his weightings. As a result, HAPLR weights Expenditures per capita with a 3 and Volumes per capita with a 1. That is, the former is treated as three times more important than the latter for the HAPLR. One can argue about whether 3:1 is reasonable: should they be unweighted? Or should the weight be 2.731:1? That kind of question is certainly a reasonable one.

Different variables vary in importance in libraries and there is no reason that fact cannot be recognized analytically. That said, there are difficulties. Statistics—the discipline—is as much art as science and arriving at reasonable judgments about measured phenomena can be tricky.

The criticism, apparently, is that since such weighting is subjective, it should not be used. Weighting is not uncommon in such analysis even with the subjective aspects and is used by the BIX. There is probably general agreement that some things we measure are more important than others. Weighting is an attempt to formalize that notion.

The third was the inclusion of input measures in the HAPLR. This criticism probably dates to a time when the quest was for outcomes from libraries. In retrospect the argument can be summarized: while outcomes were a bit vague, but good, inputs were not so useful. However: outcomes are better than outputs which are better than inputs.

What do those notions mean? We spend money on libraries—the inputs. We get outputs such as circulations. And? What do we get from circulations? Are members of the resident population better off? Do they get jobs? Lose weight? What is the outcome of all of these expenditures? How would we measure the impact of libraries on their users? How would we measure the mental units of that impact?

In the real world, one settles down and does what one can with what one has. Inputs seem to be something we can measure and financial figures are even audited. Some outputs? Yes. Outcomes? Who knows? The most commonly available data were not designed for measuring outcomes and perhaps these outcomes, like many aspects of libraries, are not even quantifiable.

Hennen's argument was that if a library has an imbalance between its inputs and outputs, it would be useful to know. Suppose we had a library with modest inputs—one with not many expenditures and few staff—which had higher outputs than a library with bountiful inputs but relatively modest outputs? Wouldn't that kind of relationship be important to know? Hence, Hennen argues, if one is trying to assess a set of libraries in a comprehensive way, one needed both inputs and outputs. He notes that the BIX included both kinds of variables and noted other similarities between the BIX's design and that of the HAPLR.

He also posits an intriguing idea. Library data are collected for several purposes but they are not designed to measure quality of a library. We all know what it is when we see it as we suggested in our discussion of the story tellers—one skilled and one less

skilled. If you stand inside the front door of a library and look around, you can usually tell if it is a good library or not by just watching what is going on. However, measuring that evanescent aspect of libraries is almost impossible because the data are not sufficiently precise or we have not yet developed methods with our data to do it. Hennen posits: perhaps imbalances of inputs and outputs may be a way to probe this question? Perhaps, indeed, if intriguing clues he discussed are corroborated with more research. Including the high and low rankings in Table 1 is a means of examining this question and is an attempt to operationalize Hennen's thinking.

Be that as it may, HAPLR published rankings based on Hennen's manipulation of the data from US government sources. He has 10 size categories for the 9,000+ libraries based on population. The reported rating numbers used varied by these size categories with a theoretical range between 1 and 1,000. Most were between 260 and 730. Libraries were given stars depending on how well they did in this scheme. HAPLR was available in *American Libraries* and via the Website.

Not all of the variables HAPLR uses are available in the Ministry's data but some are. Table 2 maps which variables in BIX, HAPLR, the LJIndex, and the US State Ranking tables (USSRT) are included in our experimental FOPL Index.

BIX

<http://www.bix-bibliotheksindex.de/en/news.html> (the English language main page)

BIX has been published for 16 years for Germany's 2,000 public and 240 academic libraries. The site is largely in German, of course, but the English pages are informative. The BIX is to be closed after the publication of the 2015 data.

There are 19 variables in the four "dimensions" used in Table 1:

Services

Usage

Efficiency

Development

<http://www.bix-bibliotheksindex.de/index.php?id=89&L=1&ref=ergebnisse>.

There is a rating scheme that, like HAPLR groups libraries by size by population served. It uses stars and a color code to rank the libraries in three percentile ranges for each of the dimensions. As a result, the effect of an unusual value within a dimension will not have an "undue influence on the overall results of the library." Damping variables with "undue influence" is a nice feature in their reporting given that we often

find errant values in library data. “BIX-TOPS” are denoted by stars. A star is awarded to libraries for each dimension in the top third and a half stars for a library for each dimension in the middle third. A library, then, can have a maximum of four stars.

The BIX is a serious effort to inventory the factors that make libraries successful according to the criteria of German libraries and to rank libraries by those factors. Our understanding of their work would benefit from a better understanding of the German terms but the organization of what they do is certainly appealing. There are dimensions of what libraries do and multiple variables in each dimension damp the effects of the occasional odd value. We might quibble here and there about this or that set of variables but the structure is sound and tested. In the case of a few of their variables, lower numbers are higher ranked. For instance, one of the measures in the Efficiency dimension is employee hours per opening hour and lower is better. But, as they point out, efficiency may well not result in good service...but, it will be cheaper for any given level of service.

Their method takes the high and low values reported and gives them a percentile. High = 100; low = 0. Then the percentiles are divided into thirds from there. As it happens, enough of the highest values we observe in our data are so high that the next value may be in the lowest third by their method. As a result, attempts to emulate their method directly ran into the problem we have with odd high or low values reported by a few libraries. It is possible that transforming the Ontario data may allow a closer emulation of this aspect of the BIX.

Given that it was a ranking scheme anyway, ranks will work, too. One thought was to divide libraries into thirds or quartiles ranked by the variables’ values but it was thought at this stage, that kind of rating might better await discussion among the membership.

BIX Dimensions are powerful tool. Given we have several variables in each dimension, we could organize variables into like groups. We decided to use these dimensions and the template, of course, can be adjusted and we can create other dimensions if we choose.

LJIndex of Public Library Service

This index is compiled by Keith Curry Lance and Ray Lyons and includes four output measures for US public libraries from the national dataset now published by the US Institute for Museum and Library Services (IMLS) and also used for HAPLR. Their goal:

“Simplicity, transparency, and comprehensibility were the main objectives in designing the rating system. We used statistical correlation analysis to identify a concise and straight-forward set of indicators of library service provision. We also wanted to enable individual libraries to examine first-hand the data upon which their ratings are based.”

<http://lj.libraryjournal.com/stars-faq/>

The four variables are based on per capita measures of:

**Library visits
Circulation
Program attendance
Public Internet computer use**

We do not have a direct measure of “public Internet computer use” in the Ministry’s data, and for visits, we have “Typical Week” data so if we take that latter figure and multiply it by 52, we will have an approximation of an annual figure. These “Typical Week” data are a bit rough—data anomalies abound—but it is what we have. They are labelled in blue in Table 1 to distinguish them.

Again, as is common, the libraries are divided by size but using Total Operating Expenditures rather than the customary population figure. As discussed above, the two sets of figures are highly correlated in these data so there is little difference between these two measures of size for our purposes and we will use the traditional resident population to group the 311 non-contracting Ontario public libraries available in the 2013 data.

It should be noted that the pairwise correlations between the three LJIndex variables in the Ontario data are all close to 1.0, the minimum being 0.97.

What is done analytically with the four variables need not concern us. It is a bit complex but the difference one gets from doing those manipulations and simply ranking the libraries by the level of the variables is negligible. In addition, like HAPLR, and the ARL Library Index the resulting numbers do not directly relate to anything. They are sums of the manipulated data.

The three LJIndex possible variables are included in the experimental FOPLIndex. The Experimental FOPL Index as seen in Table 1 merely uses ranks of 16 variables and sums those ranks across the four dimensions and for each of the size cohorts.

As a test, a comparison was made between simple ranks and ranks arrived at by emulating the calculations used in the LJIndex for the overlapping three variables. The lowest correlation was .97. In other words, simple ranks work about as well as the LJIndex's rankings by their calculated variables.

Why wasn't the correlation perfect--that is 1.0? It may be that there is an unintentional weighting fact from the method used to create the Index numbers. That method appears to be sensitive to the variance of each variable. Investigation of the details of that matter are not of central concern now. We have emulated the LJIndex by including the variables for those interested but with simple ranks.

An Experimental FOPL Index

What we seek is a useful and practical means for working librarians to assess their libraries. What can we piece together from similar efforts?

From the standpoint of mathematical rigor model of Ontario's libraries, a Factor Analysis is likely the correct approach. Such an analysis would result in a sparse index which would probably be of more use to academics than working librarians. FOPL would help but a working librarian might well be left to explain to her county or city board numbers which have no obvious meanings. We could create star awards and such but if a board member asked: how do we get our library from an index number of 1150 to say 1250? What do these kinds of numbers mean? How do they help a librarian understand the world her library is in? How does she improve the number? How does the funding agency know where the strengths and weaknesses of a library are? Our Experimental Index uses rankings, thus, a library's place in the Index is relative to other libraries, not an absolute number resulting from calculations.

What we have attempted to do is combine the best elements from previous efforts. Our Index owes most of its design to the BIX and HAPLR designs. As discussed, we follow the notion of BIX in creating "dimensions" of library activity and borrowed theirs directly. As Table 2 shows, most of the variables in this FOPL Index follow BIX's but, in fact, most of these are fairly commonly used in measuring public libraries. Table 2 shows the various ratings/indexes which use the same variable or one close.

By "close" we mean that the same variables are used but perhaps in a different way. For instance, two of these BIX variables have been altered slightly to conform to the LJIndex's related variables. The reader will recall that the LJIndex has four variables in its Public Library Service Index, one of which we do not have in the Ministry's data. BIX measures Programs per capita while the LJIndex measures program attendance per

capita. We use the LJIndex variable—with yet another high correlation (0.97)—to the BIX variable. That means, practically, the results of using the BIX for this set of variables and LJIndex's slightly different ones are about the same thing. Obviously, we gain possible comparisons with the LJIndex and lose little.

We did not ring out the last ounce of information in this design but rather, left it open for members to criticize and make suggestions. There are several paths forward with this robust template.

This tentative index is unweighted and is based on the average of ranks across all 16 variables in the five sized-based cohorts. At the bottom of the table, the reader will note other variables which could have been used. It is a mixed bag: a few have no similar data in the Ontario data. Moreover, the authors have examined public library data from two other Canadian provinces, the US, New Zealand, and Australia and have published an article based on this survey with an Australian colleague. There are many interesting data collection efforts going on which might yield useful ideas to add to any index FOPL members might wish to construct.

What we do here is try to pick a judicious mixture of variables and use unweighted average rankings and publish it with the hope that we will have comments which will help us create a better index as newer data are published.

SECTION 3. THE DETAIL

Table 3 shows the details of the formulas followed for this first FOPL Index using the Ministry's data. If you download the csv files, it is hoped that you will be able to use the documentation in this table to replicate our work. It is our hope that some will help by checking our work. Download the 2013 data here:

<https://www.ontario.ca/data/ontario-public-library-statistics>

Fair warning: emulation may be a bit complex but we welcome others joining into what we hope is a shared enterprise.

Table 3 is the result of considerable work with other spreadsheets which are summarized. These tables bridge the gap between the formulas in Table 3 and the raw data. For anyone who wishes to check our work, we will be happy to make these tables available. There is one for each variable in our experimental index, as well as one for each of the four dimensions.

Tables

Table 1	Experimental FOPL Index	5 pages
Table 1R	Experimental FOPL Index with tentative rankis	5 pages
Table 2	Variables Used in the Experimental Index	3 pages
Table 3	Computation Formulas for the Experimental Index	3 pages

Table 1: Experimental FOPL Index Results
By size cohorts and Bix-like scores

Library Number	Library Full Name	Average Rank		Resident Population		Service Dimension				Usage Dimension				Efficiency Dimension				Development Dimension					
		Lowest Rank	Highest Rank	Population	Cohort	Collection Units	Staff Population	Program Attendance	Workstations / Population	% cardholders / Population	Estimated Visits	Circulations	Stock as a %	Hours Open	Service points	Acquisitions \$ per circulation	Employees	Staff hours per open hour	Visits per \$ cost	Expenditures per employee	Staff training \$ per employee	Expenditures per \$ cost	
Cohort 1: 27 libraries with greater than 100,000 population served																							
L0082	Alex	18.4	5	27	1	119,800	24	18	27	22	23	23	24	19	15	10	10	14	17	5	25		
L0083	Beauregard	15.3	1	27	1	144,750	23	23	22	24	13	16	27	27	23	26	6	6	27	1	12		
L0084	Beauregard	15.9	1	27	1	357,800	27	26	16	23	14	1	17	14	1	5	20	6	11	1	26		
L0087	Burlington	8.3	3	16	1	174,431	3	11	7	7	7	12	7	8	11	12	15	7	3	3	4		
L0101	Cambridge	13.8	3	27	1	132,400	21	4	15	15	16	15	16	18	8	16	7	3	27	19			
L0104	Chatham-Kent	18.5	2	27	1	104,075	18	15	9	27	23	21	25	5	25	6	24	21	17	22			
L0153	Essex County	9.0	1	27	1	176,842	26	27	16	20	27	27	6	10	24	1	27	24	22	27			
L0155	Essex County	10.0	2	28	1	176,842	13	13	6	3	4	24	2	9	20	2	23	24	22	27			
L0189	Greater Sudbury	16.9	1	14	1	182,230	12	46	24	4	4	2	1	4	4	10	27	9	9	7			
L0214	Hamilton	6.1	2	17	1	540,000	26	8	8	19	5	4	2	1	4	10	27	9	9	7			
L0213	Kingsville-Frontenac County	11.8	2	22	1	151,650	10	13	4	2	13	17	20	5	6	19	9	22	13	2	13		
L0127	Kitchener	14.6	8	24	1	284,000	8	12	23	11	15	14	19	24	13	23	8	11	8	18			
L0239	Lambton County	12.0	1	27	1	184,100	5	7	2	2	25	26	1	1	27	7	26	10	13	10			
L0245	London	8.2	1	22	1	372,730	6	1	1	10	12	7	9	13	12	11	5	4	4	10			
L0099	Markham	11.0	1	25	1	332,889	17	14	6	20	3	4	22	25	17	19	1	1	1	4			
L0098	Mississauga	15.4	2	25	1	754,000	20	22	17	21	10	19	8	18	4	13	4	2	20	24			
L0247	Oakville	11.1	3	25	1	185,100	4	14	4	4	11	9	12	16	5	8	13	12	6	9			
L0253	Oshawa	13.8	3	26	1	157,000	18	4	14	21	8	10	15	7	7	24	19	26	12	6			
L0481	Ottawa	14.4	9	21	1	842,238	19	10	12	20	16	9	11	14	11	9	14	16	15	21			
L0294	Richmond Hill	14.2	2	26	1	194,335	3	19	16	13	16	14	6	18	6	26	2	16	14	10			
L0024	St. Catharines	16.7	5	25	1	131,400	12	7	25	19	5	25	20	19	21	17	22	25	18	23			
L0349	Thunder Bay	14.0	5	22	1	108,359	8	9	13	5	17	22	13	15	21	15	17	22	16	6			
L0353	Toronto	11.6	2	23	1	2,711,770	6	5	10	13	6	10	17	10	12	14	17	12	19	23			
L0359	Vaughan	17.5	5	26	1	315,006	24	16	5	26	9	19	18	6	25	23	2	2	2	2			
L0385	Waterloo City	12.1	2	23	1	131,266	11	12	17	11	6	12	2	5	23	2	20	3	5	19			
L0277	Whitby	13.8	3	23	1	131,360	23	17	3	17	16	5	3	20	21	3	18	10	14	9			
L0278	Windsor	17.3	2	26	1	210,831	15	21	2	25	24	21	8	3	18	4	4	25	25	15			

**Table 2: Variables used in Experimental FOPL Index
With a comparison to BIX / HAPLR / LJIndex**

Service Dimension	Variables in Experimental FOPL Index	Same or related variables also used in:
	Collection Units per capita	BIX / HAPLR / USSRT
	Population Served / Staff	HAPLR
	Program Attendance per capita	BIX / LJIndex
	Population Served / Workstations	BIX / USSRT
	Population Served / Registered Borrowers	USSRT
Usage Dimension	(Estimated) Visits per capita	BIX / HAPLR / LJIndex / USSRT "Typical Week" data
	Circulations per capita	BIX / HAPLR / LJIndex / USSRT
	Turnover as a %	BIX / HAPLR
	Hours open per capita	BIX
	Resident Population / Service points	BIX
Efficiency Dimension	Collection Expenditures per circulation *	BIX / HAPLR Smaller is better
	Staff hours / open hours *	BIX Smaller is better
	(Estimated) Visits per open hours	BIX / HAPLR "Typical Week" data
	Total Operating Expenditures per (Estimated) Vis	BIX Smaller is better "Typical Week" data

Total Operating Expenditures per capita

HAPLR / USSRT

BIX not used

User floor space in sqm per 1000 capita

Employees per 1000 capita

Internet services

Virtual Library Visits per capita

Renewal rate

Capital investment per capita

No measures of size of libraries

Not clear what this means

Too many missing observations

No data

Fluctuates too much

HAPLR not used

Percent Budget to materials

Materials Expend. Per capita

Periodicals per 1000 residents

Circulation per FTE Staff Hour

Reference per capita

Circulation per hour

Circulation per visit

No similar measure

LJIndex not used

Public Internet computer use

No similar measure

US State Ranking Tables (FY 2012) not used

- Number of interlibrary loans received per capita
- Average number Internet computers per stationar *"per service point"* in BIX and Ministry data terminology
- Current print serial subscriptions per capita
- Number of audio-physical units per capita
- Number of video-physical units per capita
- Number of paid full-time-equivalent (FTE) staff per capita
- Number of paid FTE librarians per capita
- Number of paid full-time-equivalent (FTE) librari; No similar measure with an "ALA-MLS" per capita
- Number of other paid FTE staff per capita
- Total operating revenue per capita
- Total state operating revenue per capita
- Local operating revenue per capita
- Other operating revenue per capita
- Total collection expenditures per capita
- Total staff expenditures per capita
- Salaries and wages expenditures per capita

Table 3: Computation Formulas for the Experimental FOPL Index
 Using Ontario's Ministry of Tourism, Culture, and Sport Library Data Spreadsheet for 2013

Formulas

Per capita = per resident population
 and so on below

Data source from Ministry spreadsheet:

P1.1 - Resident Population Served

Service Dimension Collection Units per capita

Sum the following figures:

- C1.2.T - Total Volumes Held (reference)
- C1.3.3.T - Total No. of copies of CD and DVDs (reference)
- C2.2.T - Total Volumes Held
- C2.3.2.T - Total No. of copies of CD and DVDs
- C2.3.4.T - Total No. of Copies of E-resources, including E-books
- C3.2.1.T - Total No. of Database Subscriptions
- C3.2.3.T - Total No. of copies of CD and DVDs
- C3.2.4.T - Total No. of individual electronic periodicals titles
 and divided by Resident Population

Population Served / Staff counts

Staff = sum of these counts

- D1.1.1.C - Librarians
 - D1.1.2.C - Library Technicians
 - D1.1.3.1.C - Other Professional Staff
 - D1.1.3.C - EXCEL Graduates
 - D1.1.4.C - Other Staff
 - D1.1.5.C - Volunteers
- (plus the sum of hours worked by part-time employees*
- D1.2.1.H - Librarians
 - D1.2.2.H - Library Technicians
 - D1.2.3.1.H - Other Professional Staff
 - D1.2.3.H - EXCEL Graduates
 - D1.2.4.H - Other Staff
 - D1.2.5.H - Volunteers
- divided by*
- D1.0.H - How many hours per week in your library constitute
 a full time position

Program Attendance per capita	<u>F2.2.A - Annual program attendance</u>
Population Served / Workstations	E1.1 - Total No. of Public Access Workstations available in your library
Population Served / Library Cardholders	<u>A1.14 - No. of Active Library Cardholders</u>
Usage Dimension (Estimated) Visits per capita	<u>G1.5.1.W - No. of visits to the library made in person</u> (these are "Typical Week" data hence: this number times 52 for an estimated annual figure)
Circulations per capita	F1.0 - Total Annual Circulation (Actual Annual Direct Circulation)
Turnover as a %	Operationalized as: collection units as a percent of Total Annual Circulation. Both are defined above
Hours open per capita	<i>Hours open = sum of:</i> E3.1.SPH - Main Library E3.2.SPH - Branches open 12 hours per week or more E3.3.SPH - Branches open less than 12 hours per week E3.4.SPH - Bookmobile stop locations E3.5.SPH - Deposit Stations and Kiosks (refer to criteria)
Resident Population / Service points	<i>Service points = sum of</i> E3.1.SPC - Main Library E3.2.SPC - Branches open 12 hours per week or more E3.3.SPC - Branches open less than 12 hours per week E3.4.SPC - Bookmobile stop locations E3.5.SPC - Deposit Stations and Kiosks (refer to criteria)
Efficiency Dimension Collection Expenditures per circulation *	<u>B4.01 - Materials (expenditures) /</u> F1.0 - Total Annual Circulation (Actual Annual Direct Circulation)
Staff hours / open hours *	<i>Staff hours = sum of (FTE):</i> (D1.1.1.C - Librarians D1.1.2.C - Library Technicians D1.1.3.1.C - Other Professional Staff D1.1.3.C - EXCEL Graduates

- D1.1.4.C - Other Staff
- D1.1.5.C - Volunteers)
times
- D1.0.H - How many hours per week in your library constitute a full time position)
(plus the sum of hours worked by part-time employees
- D1.2.1.H - Librarians
- D1.2.2.H - Library Technicians
- D1.2.3.1.H - Other Professional Staff
- D1.2.3.H - EXCEL Graduates
- D1.2.4.H - Other Staff
- D1.2.5.H - Volunteers
divided by the sum of the hours open figures:
- E3.1.SPH - Main Library
- E3.2.SPH - Branches open 12 hours per week or more
- E3.3.SPH - Branches open less than 12 hours per week
- E3.4.SPH - Bookmobile stop locations
- E3.5.SPH - Deposit Stations and Kiosks (refer to criteria)

(Estimated) Visits per open hours

G1.5.1.W - No. of visits to the library made in person
Given these are "Typical Week" data and open hours are also, this figure and the open hours figure from above are used.

Total Operating Expenditures per (Estimate) B5.0 - Total Operating Expenditures

divided by

G1.5.1.W - No. of visits to the library made in person
(these are "Typical Week" data hence: this number times 52 for an estimated annual figure

Development Dime Staff training \$ per employee

B4.03 - Staff Training (Total funds spent on staff training, including travel and accommodation)
divided by
Staff counts (from above)

Total Operating Expenditures per capita B5.0 - Total Operating Expenditures