

Recent developments in library evaluation, statistics and measurement – and why they are important

Sharon Markless

King's Learning Institute, King's College, London, UK.
sharon.markless@kcl.ac.uk

David Streatfield

Information Management Associates, Twickenham, UK.
streatfield@blueyonder.co.uk



Copyright © 2014 by Sharon Markless and David Streatfield. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License:
<http://creativecommons.org/licenses/by/3.0/>

Abstract:

This paper identifies six key changes in library evaluation over the past five years and suggests why they may be important for IFLA delegates. Deciding which changes to feature was difficult because library evaluation is a highly contentious area (as is evaluation more generally) and because practice across the world is partial, patchy and uneven. The authors summarize their choices by noting a gradual move in the library evaluation field towards deeper understanding of the nature and significance of evaluation. Their key propositions are that: the demand for evaluation expertise is beginning to outstrip supply; there is a technology-enabled shift in research and evaluation focus from the world of the library to the information world of potential library users; there is growing experience of conducting library impact evaluation across countries; there is a dawning recognition that sustaining impact evaluation efforts is a core element in ensuring sustainability of libraries; there are early signs of movement towards more ethical evaluation; and as library services become more innovative the limitations of simple evaluation models are beginning to show. The implications of each of these propositions for IFLA delegates is adumbrated.

Keywords: Impact evaluation; library statistics; library measurement; library evaluation; sustainability of libraries.

Introduction

When we were invited to make this presentation some months ago, we were asked to identify five things that have happened in library evaluation in the past five years and why they are important for the work of people at this conference. Trying to decide what constitutes important ‘things that have happened’ was not straightforward for several reasons, such as:

Library evaluation is still a highly contentious area, with little agreement between proponents of different evaluation ‘schools’. In this respect, library evaluators mirror many of the wider evaluation debates and discussions in disciplines such as education or international development evaluation. There, as in the library evaluation field, some evaluators define their evidence collection activity entirely or very largely in quantitative terms; others ‘restrict’ themselves to qualitative evaluation (see, for example, Guba and Lincoln, 1987; Shaw, Greene and Mark, 2006); still others opt for mixed methods approaches, which they may view as relatively unproblematic or as intensely complex and difficult (see Patton, 2012; in the library context see Chirarov, 2012). We will be disappointed if everyone agrees with our perceptions of what the key changes in the field are and why they are important (we don’t always agree between ourselves).

Practice in library evaluation varies in different countries and the overall picture is partial, patchy and uneven. If we think about conducting performance measurement and library evaluation at country level (which seems appropriate for an international federation such as IFLA) we need to ask ourselves where in the world we are thinking about. Library evaluation can be seen as alive and well – if we focus on the leading work, which tends to be conducted in the USA and Canada, Australasia and Northern Europe (see Markless and Streatfield, 2013; Poll, 2014). The library evaluation situation in many other countries is much less rosy. This makes any choice of significant developments open to the question ‘Significant for whom?’

We could have highlighted changes in how to do impact evaluation, what to evaluate or particular techniques, such as:

- the new International Standard *Methods and Procedures for Assessing the Impact of Libraries*, ISO 16439
- various return on investment studies – usually based on contingent valuation, with its inherent limitation of being based on hypothetical situations. A valuable exception here is the work of Tenopir and her colleagues focused on scholarly reading and the value of library resources (Tenopir et al. 2010; 2012)
- using collaborative performance data collection and analysis tools (see, for example, Hall, Stephens and Kennedy, 2014; Killick, van Weerden and van Weerden, 2014).

Instead, we decided to focus on what we saw as recent changes in thinking about evaluation – which could be summarised as a gradual move in the library evaluation field towards deeper understanding of the nature and significance of evaluation – or more flippantly as moving beyond collecting and analysing what we refer to as ‘busyness statistics’ (service process and output indicators) towards engagement with user and potential user perceptions of libraries. We were bolstered in our decision by the knowledge that the other speakers in this session will explore specific, concrete developments. Taking this view of the evaluation scene, we almost succeeded in meeting the brief (we have identified six important changes rather than five, but hope to be forgiven for this lapse).

A final preliminary point (if that is not a contradiction in terms) is that, in thinking about important changes, we have drawn on our experience of working with the Global Libraries Initiative of the Bill and Melinda Gates Foundation at several points. This is because we see that Initiative as driving some recent developments that are likely to resonate with many IFLA delegates.

What do we see as changing?

1 The demand for evaluation expertise is beginning to outstrip supply

Doing performance measurement and evaluation in libraries requires a range of skills which become more disparate as we move beyond ‘business statistics’ towards user-focused evaluation.

Our experience working as independent impact consultants to the Global Libraries Initiative highlights what we see as a growing issue. Global Libraries has awarded a series of country grants in emergent countriesⁱ which, *inter alia*, require the country teams to procure hardware and software to provide public library users with free public access to computers and the internet, train library staff to support users in computer and internet use, and encourage library managers to devise advocacy programmes to help sustain this work. (For a fuller description of the Global Libraries initiative and of its approach to impact evaluation see Fried, Kochanowicz and Chiranov, 2010). They ask that each country team should include an impact specialist to ensure that appropriate performance measurement and impact evidence gathering is carried out to show whether and how providing free public access to computers is changing people’s lives. Many of the countries involved have found it difficult to find suitable impact specialists locally and none of the impact specialists appointed has had the full range of requisite skills at the outset. This is scarcely surprising because this role calls for a combination of expertise in quantitative and qualitative evaluation in national development programme settings, as well as understanding of how public libraries and their staff function in the relevant country. Impact specialists have been recruited from a variety of backgrounds: international development NGOs, business management and consultancy, health strategy evaluation or rural development. Part of our role with Global Libraries has been to help people with this range of backgrounds to get to grips with public libraries.

This ‘shortfall in expertise’ is not peculiar to Global Libraries, nor is it in any way surprising. When library policy makers and service managers focus their attention beyond service performance measurement and towards the impact of their services on users, they often encounter similar issues. Participants in our impact workshops for library managers reported, over several years, that library policy makers were often disconcerted when presented with qualitative evidence of the impact of library services on people rather than the tabulated performance metrics with which they felt comfortable. In turn, the library managers usually found it challenging to work through a training process which required them to clearly state service objectives and then to identify relevant impact indicators, before choosing appropriate evidence collection methods.

1.1 What are the implications for IFLA delegates?

Library schools and Library Associations should think about how to develop a wider range of evaluation skills amongst the library communities within their countries. (We will return to this theme at 4 below.)

2 There is a technology-enabled shift in research and evaluation focus from the world of the library to the information world of potential library users

We detect a gradual change in focus from a library-centred view of evaluation (concentrating on performance metrics and user satisfaction surveys), towards studying ‘the metrics of everyday professional life’ (looking at people’s information worlds and whether and where libraries and their services might fit in). Evaluators and researchers are enlisting the power of computers to gather and process evaluation evidence through a growing variety of means, from pop-up surveys (see Chiranov, 2011) through to electronic tracking of people’s information search behaviour (or ‘digital footsteps’) (Nicholas and Rowlands, 2008). These approaches enable evaluators to move on from reliance on how people say that they use information/libraries in their work, education and social lives to seeing how they actually use information, including library use.

The rapid development of mobile technologies has enabled libraries to be embedded into people’s work patterns (seen by some commentators as a key step towards library survival). This provides a further challenge in evaluating the reach of such libraries.

This process of technology enablement calling for more sophisticated evaluation will continue - as information technologies develop, they can provide interesting and valuable ways of engaging with library evaluation, especially changes in how users interact with information.

2.1 What are the implications for IFLA delegates?

There is an opportunity for you to consider:

- adapting traditional methods of data collection and asking different questions
- using information technologies in different ways – e.g. observatories drawing information search data from linked computers
- the ethical issues (evaluation versus spying)

3 There is growing experience of conducting library impact evaluation across countries

During our seven-year experience of working with the Global Libraries initiative we have seen a steadily growing interest by participants in being able to compare performance measurement data and impact evidence between countries and across the whole programme. This movement has been accompanied by a shift in focus from process and output measurement towards outcomes. Making comparisons is an activity fraught with difficulties at any level (institutional, regional or country). The evaluator is open to the charge of a lack of contextual sensitivity and therefore drawing invidious conclusions. Comparison has to be both systematic and rigorous, particularly at this large-scale level. Part of the contribution that Global Libraries should be able to make to the world’s library evaluation community is a well-tried and validated system for gathering common impact measures at country level using common survey questions and based on well-researched impact areas (where we know that public libraries are capable of making a difference). Global Libraries has already made available the survey tools and other instruments used in their European Community-focused 17 country survey of public perceptions of public libraries (Quick *et al.*, 2013). These comprise survey instruments for a general population survey, computer users in public libraries, computer and library users, library users; an interview instrument for library managers; a library users focus group instrument; and a qualitative research stimulus guide (as well as country reports for each country involved in both English and the main country

language and, importantly for other evaluators, access to the raw data). All these can be found on the TASCHA website at <http://tascha.uw.edu/research/libraries/publications/>.

Again, this groundswell of change is not confined to Global Libraries. To give two examples, EIFL blazed the trail for the work just described by commissioning a survey of perceptions of public libraries in six African countries (TNS RMS East Africa, 2011) and has again made the survey instruments and tools accessible for the library evaluation community at <http://www.eifl.net/perception-study>. IFLA itself has taken a lead in encouraging impact-focused evaluations through at least two of its initiatives, FAIFE (*Committee on Freedom of Access to Information and Freedom of Expression*) and BSLA (Building Strong Library Associations), in the latter case actively encouraging the countries involved to conduct their own impact evaluations and opening up the possibility for other countries to apply similar approaches to expand the base for comparison across countries.

3.1 What are the implications for IFLA delegates?

If you are considering country-level library surveys focused on people's perceptions or impact, it should at least be worthwhile to look at some of the instruments that are already available. They offer a potentially enhanced basis for comparisons with other countries founded on more rigour in common sample sizes, explicit assumptions, research-based question areas and common questions. This should in turn provide opportunities to make the right (well-informed) comparisons for the right (developmental rather than judgemental) comparisons.

If library policy makers and managers are looking to gather comparative data within their counties they might also find these instruments and the underlying approaches provide useful starting points.

4 There is a dawning recognition that sustaining impact evaluation efforts is a core element in ensuring sustainability of libraries

We hope that we are not being over-optimistic in detecting a growth in awareness of the importance of performance measurement and impact evaluation data as advocacy evidence. If libraries are to survive as one of the plethora of information-access agencies they need to adapt and respond to changing user needs. This calls for sustained user impact evaluation efforts to be encompassed within normal work practices of library managers, as well as when planning how to continue effective library initiatives after time-limited funding has disappeared.

Successfully sustaining evaluation depends on empowerment of library managers as evaluators, based on helping them to understand what is involved. In the UK the way forward was initially through individual self-empowerment based on training, supported by access to appropriate toolkits and guidance. However, various early approaches to supported self-evaluation (Markless and Streatfield, 2004; 2006) are now being transformed and developed. For example, in Romania a voluntary Impact Group of methodologists and library managers is developing evaluation tools and has been trained to carry on elements of impact evaluation after conclusion of the Global Libraries grant there later this year. In Lithuania, the *Libraries for Innovation 2* programme is encouraging library managers to take on an element of impact evidence gathering and analysis, again with training provided, as part of 30 small grants to develop public library services.

4.1 What are the implications for IFLA delegates?

We obviously think that supported self-evaluation is potentially important in many countries if we are to address the skills gap referred to earlier. However, we also recognise the need to be realistic – how far is supported self-evaluation possible in environments where the reality may be that libraries are staffed by middle-aged or older people who are employed part-time, may be relatively poorly educated and under-motivated to accept any challenges in their work?

What can library policy makers and managers do to help? It should be possible to ensure that appropriate toolkits and guidance are available and accessible, and to encourage and support library managers who want to do this work (for example, by building library manager support into library evaluation strategies). We also recommend that sustainability of evaluation should be seen as part of any overall library sustainability strategy. The future of libraries depends upon library policy makers and managers accepting the need for evidence-informed working in order to understand the accelerating rate of change that will constitute a big part of all our futures.

5 There are early signs of movement towards more ethical evaluation

Libraries are getting better at engaging with communities in deciding what services to develop and how to develop them, but evaluation still tends to be controlled by professional evaluators or library managers. However, we are beginning to detect some movement away from treating library users as ‘evaluation subjects’ and towards more democratic evaluation. The international development evaluation literature has advanced various ideas about a more inclusive and democratic evaluation embracing a wider range of stakeholders, including marginalized groups, and involving them in designing the evaluations and in interpreting the findings (e.g. Mertens, 2003; Greene, 2006). Central here is the idea of preventing qualitative evaluation from becoming just another way of enforcing the existing power relationships between governments and their people. In library evaluation this idea takes various forms, ranging from consultation with users about the impact areas for evaluations and forming evaluation partnerships with stakeholders to establishing panels of library staff and users to review impact evidence and take editorial decisions about the presentation of results.

5.1 What are the implications for IFLA delegates?

A good starting point may be to consider three questions: Who are the evaluations for? Who should decide what to evaluate and how? Who owns the results and decides what to do with them?

Going a step further, in an examination of the consequences of 23 research projects funded by Canada’s International Development Research Centre, Fred Carden (2009) identified three ways in which research can contribute to better governance, as:

- 1 Encouraging open inquiry and debate
- 2 Empowering people with the knowledge to hold governments accountable
- 3 Enlarging the array of policy options

These three principles appear to present an appropriate challenge for library service evaluators.

6 As library services become more innovative the limitations of simple evaluation models are beginning to show

The usual approach to library evaluation is sometimes called the simple logic model and is based on assuming that there is a series of direct relationships between service inputs, outputs and outcomes (expressed as quantified outputs). However, as libraries increasingly seek to take advantage of ICT developments by introducing new services, design different spaces to engage the community in new activities and venture into teaching and learning in the area of information literacy, they are moving into the area of ‘emergent evaluation’ (sometimes described as ‘developmental evaluation’ – Patton, 2011) in which it is not straightforward to predict where change will appear and therefore, where to focus the evaluation. Rogers (2008) distinguishes three types of model. At the two extremes are the simple logic model and the complex logic models where the paths from action to impact are complex, with disproportionate relationships (where, at critical levels, a small change can make a big difference) and emergent impacts (which cannot readily be specified at the outset). The more advanced library services appear to have all the characteristics of the complex logic model, but so far we have seen more discussion than action in addressing the issue here.

6.1 What are the implications for IFLA delegates?

When engaging in the more complex areas of library service provision it may be necessary to adopt a more flexible approach to evaluation which does not assume that we can guess all the likely outcomes from a programme or initiative at the outset. The idea of progressive focusing by collecting as much evaluation evidence as possible and then gradually deciding which impact areas are likely to repay greater attention is not new (at least in educational evaluation) but may offer a helpful way forward (see Parlett and Hamilton, 1972).

Most library statistics are based on the simple logic model and their interpretation is often accordingly simplistic. Is it possible to use library statistics (in combination with other evaluation data) in ways that are sensitive to how change occurs in complex library environments? We don’t know the answer to this question so it is probably a good place to stop looking over our shoulders.

What will the next five years bring to the evaluation party? Hopefully, library evaluators will engage more with their users to explore some of the intangible but important aspects of the library contribution to people’s lives. The best library evaluators are already looking at user confidence when assessing ICT-focused training, rather than stopping at competence. We still have something to learn from our cultural sector colleagues when exploring creativity and enjoyment and how these can be enhanced as part of the library experience.

DISCLAIMER

We have made various references above to the Global Libraries initiative of the Bill and Melinda Gates Foundation. We wish to emphasise that these comments are *our* views and should not be taken in any way to represent the official view of Global Libraries or of the Bill and Melinda Gates Foundation.

References

Carden, F. (2009) *Knowledge to Policy: making the most of development research*, Sage/IDRC, Thousand Oaks, CA, available at www.idrc.ca/openebooks/417-8/

Chiranov, M. (2011) Applying Pop-up Survey Software to Incorporate Users' Feedback into Public Library Computing Service Management, *Performance Measurement and Metrics*, **12** (1), 50–65.

Chiranov, M. (2012) Combining Qualitative and Quantitative Evidence: forming strategic collaborations – and telling stories. In Streatfield, D. R., Paberza, K., Lipeikaite, U., Chiranov, M., Devetakova, L. and Sadunisvili, R. (2012) Developing Impact Planning and Assessment at National Level: addressing some issues, *Performance Measurement and Metrics*, **13** (1), 63–5.

Fried, S., Kochanowicz, M. and Chiranov, M. (2010) Planning for Impact, Assessing for Sustainability, *Performance Measurement and Metrics*, **11** (1), 56–74.

Greene, J. C. (2006), Evaluation, Democracy and Social Change. In Shaw, I.F., Greene, J.C. and Mark, M.M. *Handbook of Evaluation: policies, programs and practices*, London, Sage.

Guba, E. G. and Lincoln, Y. (1987) *Fourth Generation Evaluation*, London, Sage.

Hall, I., Stephens, J. and Kennedy, S. (2014) [Can you measure IT? The UK Experience of TechQual+](http://www.emeraldinsight.com/journals.htm?issn=1467-8047&volume=15&issue=1), *Performance Measurement and Metrics*, **14** (1/2), <http://www.emeraldinsight.com/journals.htm?issn=1467-8047&volume=15&issue=1>

Killick, S., van Weerden, A. and van Weerden, F. (2014) [Using LibQUAL+® to Identify Commonalities in Customer Satisfaction: the secret to success?](http://www.emeraldinsight.com/journals.htm?issn=1467-8047&volume=15&issue=1), *Performance Measurement and Metrics*, **1/2**, <http://www.emeraldinsight.com/journals.htm?issn=1467-8047&volume=15&issue=1>

Markless, S. and Streatfield, D. R. (2004) *Improve your Library: a self evaluation process for secondary school libraries and learning resource centres*, 2 vols, London, Department for Education and Skills, www.informat.org/ (click on School Libraries) .

Markless, S. and Streatfield, D. R. (2006) Gathering and Applying Evidence of the Impact of UK University Libraries on Student Learning and Research: a facilitated action research approach, *International Journal of Information Management*, **26**, 3–15.

Markless, S. and Streatfield, D.R. (2013) *Evaluating the Impact of Your Library*, London, Facet Publishing.

Mertens, D.M. (2003) The Inclusive View of Evaluation: implications of transformative theory for evaluation. In Donaldson, S.I. and Scrivens, M. (eds) *Evaluating Social Programs and Problems: visions for the new millennium*, Mahwah, NJ, Lawrence Erlbaum Associates.

Nicholas, D. and Rowlands, I. (2008) *Digital Consumers: reshaping the information professions*, London, Facet Publishing.

Parlett, M. and Hamilton, D. (1972) Evaluation as Illumination: a new approach to the study of innovatory programmes. Republished in Parlett, M. and Dearden, G. (eds) (1977) *Introduction to Illuminative Evaluation: studies in higher education*, Cardiff-by-the-Sea, CA, Pacific Soundings Press, and Guildford, Society for Research into Higher Education.

Patton, M. Q. (2011) *Developmental Evaluation: applying complexity concepts to enhance innovation and use*, New York, NY, The Guilford Press.

Patton, M. Q. (2012) *Essentials of Utilization-focused Evaluation*, Thousand Oaks, CA, Sage.

Poll, R. (2014) *Bibliography: Impact and Outcome of Libraries*, The Hague, Netherlands, IFLA.

Quick, S., Prior, G., Toombs, B., Taylor, L. and Currenti, R. (2013) *Cross-European Survey to Measure Users' Perceptions of the Benefits of ICT in Public Libraries* (funded by the Bill & Melinda Gates Foundation), <http://www.ifla.org/publications/publications-associated-with-the-s-e-sectionhttp://tascha.uw.edu/publications/cross-european-survey-to-measure-users-perceptions-of-the-benefits-of-ict-in-public-libraries/>

Rogers, P. J. (2008) Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions, *Evaluation*, **14** (1), 29–48, <http://evi.sagepub.com/cgi/content/abstract/14/1/29>.

Shaw, I. F., Greene, J. C. and Mark, M. M. (2006) *Handbook of Evaluation: policies, programs and practices*, London, Sage.

Tenopir, C. and Volentine, R. with King, D.W. (2012) *UK Scholarly Reading and the Value of Library Resources*, JISC Collections, www.jisc-collections.ac.uk/news/UK-scholarly-reading

Tenopir, C., King, D., Mays, R., Wu, L. and Baer, A. (2010) Measuring Value and Return on Investment of Academic Libraries, *Serials*, **23** (3), 182–90.

TNS RMS East Africa (2011) *Perceptions of Public Libraries in Africa: combined report*, http://www.eifl.net/system/files/201108/perceptions_study_coverpage.jpg

ⁱ In addition to work in the United States, Global Libraries country grants have supported efforts in Mexico, Chile, Colombia, Jamaica, Botswana, Lithuania, Latvia, Romania, Ukraine, Poland, Bulgaria, Moldova, Turkey, Vietnam, and Indonesia.