



Increasing Library Intelligence by Harmonizing Assessment Tools¹

Aggeliki Giannopoulou

Library & Information Center, University of Patras, Greece
agianop@upatras.gr, orcid.org/0000-0001-7221-7055

Georgia Giannopoulou

Library & Information Center, University of Patras, Greece
ggiannop@upatras.gr, orcid.org/0000-0001-7703-846X

Alexandra Goudis

Library & Information Center, University of Patras, Greece
agoudi@upatras.gr, orcid.org/0000-0002-9329-3249

Giannis Tsakonas

Library & Information Center, University of Patras, Greece
gtsak@upatras.gr, orcid.org/0000-0002-8786-9440

Abstract

Academic libraries are considered as key factors in the educational system of a country and strong pylons for the economic development and societal cohesion. Libraries have always intended to provide qualitative services and frequently run surveys that measure the opinion of their users. Our contribution aims to show how the findings of a survey can be aligned with information from other assessment tools to better inform library management. We analyze external data of our Library's performance as collected by an electronically conducted survey in May 2018. The initial objective was

to collect 1,000 questionnaires from all registered Library users; however, 950 questionnaires were collected in a two weeks' period, securing the quota sampling conditions. Descriptive statistical analysis via SPSS was conducted to find the key measurements and to explore deeper the various associations. At the same time, we compared the scores of library performance as reflected in internal managerial assessment practices over the span of two years. This study follows an approach that exploits the existing percentage measurement scale for the assessment of Greek public sector employees to gather users' opinion on certain performance categories. This scale, as well as its interpretation module, is used for the harmonization of the information that comes from varied assessment notions, tools and practices. Our survey findings showed that the library users were well satisfied with the conduct of staff and think that the library has margins for improvement to fulfil its role as a study place and collection. These findings seem close to the respective internal assessment scores. Therefore, our study showcases that there can be a harmonization of various assessments tools, internal and external, and the library administration can be informed about its performance in a coherent way.

Keywords: library assessment; staff performance; user survey; library intelligence; decision making

1. Introduction

Libraries have a long standing tradition of evaluation and reporting. Their commitment to quality services and the constantly challenging environment require library personnel to assess regularly their performance on a wide range of criteria. Quite often assessments are being conducted by different organizations, including the public administration bodies that oversee the operation of libraries. However, one of the effects of this heterogeneity is that most of these initiatives use instruments and scales that do not relate one to another and that there is no unambiguous way of interpreting their results. The differences in the semantics of criteria, as well as in the scales, such as the intervals they use, are impediments to the unified understanding of the performance of a library. Furthermore, distinct, asynchronous and non-associated evaluation practices increase demands in resources and in many institutions it is difficult to become part of the regular workflow (Dole, Liebst, & Hurych, 2006). In our contribution we were motivated by the need to have a light-weight evaluation instrument that is not constrained by

domain-specific elements and is connected with the Public Sector practices for performance evaluation. Therefore, we exhibit the rationale of the design of our instrument and we show how the findings of a user survey were aligned with information from other assessment tools in order to better inform library management.

2. Background

The library community has tried to solve the issue of heterogeneity through the adoption of ISO standards, such as ISOs 2789 (ISO, 2013) for output, 11620 (ISO, 2014a) for outcomes and 16439 (ISO, 2014b) for impact assessment; yet their application is not globally spread. Renard (2007) mentions as potential reasons the lack of a solid, common perception of the purpose of the Standards, as well as the time lag between the standardization processes and the rapidly evolving library environment. Other similar initiatives include community tools and protocols, such as LibQUAL+® (Association of Research Libraries, n.d.) or MINES for Libraries® (Association of Research Libraries, n.d.). Bertot (2001) underlined the need to focus on internationally aligned initiatives for the comparison of longitudinal performance measures and although this stands true for the electronic information, in the area of community satisfaction and perceived performance this requires another approach that enhances the contextualization of the evaluation. Franklin, Plum and Kyrillidou (2009), while taking into account the various challenges of the networked environment, mention four dimensions for e-metrics evaluation, namely (a) externally generated, vendor usage data, (b) locally or internally generated data, (c) externally generated, web survey data and (d) internally generated, web survey usage data. These dimensions are indicative of what is the affordance and what are the constraints of each source of data, such as whether it is internally or externally administered and the degree of representation.

In Greece there are several notable evaluation initiatives on the institutional or national level, including the statistics for the library sector of the Statistical Agency of Greece (Greek Statistical Authority, 2019) or the performance measurement indicators of the Quality Assurance Unit of HEAL-Link (HEAL-Link, 2019), the consortium of Hellenic academic libraries. These concentrate mostly on system-centered indicators, such as figures about

collection development, circulation, ILL and so on. User surveys are often using different criteria and scales. This has been noted also in the institutional-wide user surveys that our Library has conducted in the past (Library & Information Center, University of Patras, 2019). Therefore, differences are not only observed between various organizations and various levels, but also within the same organization in different time-frames.

In most European countries, academic libraries are considered part of the public sector. Despite the fact that they operate in the academic environment, which in principle gives them autonomy and attaches them to the vision and the mission of their institution, they still remain organizations of the public sector and their employees are considered public servants. As Richard mentions “There is almost no internal (library or institutional) pressure independent of government” (Richard, 1992) and libraries are not unaffected by the governmental decisions, while at the same time, from the viewpoint of public administration, evaluation should be considered as “a mean which helps libraries to arrange their activities towards high-quality information services for the users” (Rudžionienė & Dvorak, 2014).

Our Library still contributes to the aforementioned national evaluation efforts, which are focusing mostly on internally generated data. The State-driven collection of statistics is in accordance to the statistics provided to the Quality Assurance Unit, but for the internally generated data through web surveys, a different approach was required. We believed that the domain specific instruments were not properly linked to State-initiated evaluation schemes. Therefore, we wanted to find common ground where we would align the evaluation instruments and this was found in adopting a scale that has been widely used in the context of the country.

3. Setting

For this study we used data from two distinct sources. First, we used anonymized scores of all personnel members in the Public Sector Questionnaire (henceforth PSQ) by the First Tier Manager for the biannual period 2016–2017 and we compared the scores. In Greece, the heated debate of the evaluation of public sector personnel was resolved in 2016 with a new scheme by the Ministry of Administrative Reform. All public servants,

including academic librarians, are now evaluated by their two upper level managers, the first and the second tier manager in ten criteria, which are divided in three categories: (a) Knowledge, Interest and Creativity, (b) Ethics and Behaviour and (c) Effectiveness (Ministry of Administrative Reform, 2016). The evaluation instrument is using a percentage scale that has a detailed interpretation schema, which describes the state of performance of each employee in these criteria (Figure 1).

Second, we used the data from the user survey that our Library conducted in May 2018. The survey lasted for two weeks and our initial objective was to have 1.000 entries from a stratified sample population. We aimed having an analogous representation of all of our library patron categories, but we preferred to avoid further constraining it by disciplines or departments. In order to engage our population, prize draws were advertised, calls for participation were repeated at regular intervals, the survey instrument was accessible online and it was designed in a way that used extensively sliders in a responsive webpage. Our instrument had the same characteristics as PSQ, i.e. a percentage scale, and our intention was to make it usable on mobile devices, such as pads and smartphones, which are very popular to the younger parts of our community. This was supported by Arnau, Thomson and Cook (2001) that the slider “gives the impression of an anchored, continuous scale, which is attractive from a psychometric perspective.” We did not though announce the interpretation of the scale to the public to let our users rate our Library

Fig. 1: The Public Sector Questionnaire scale.

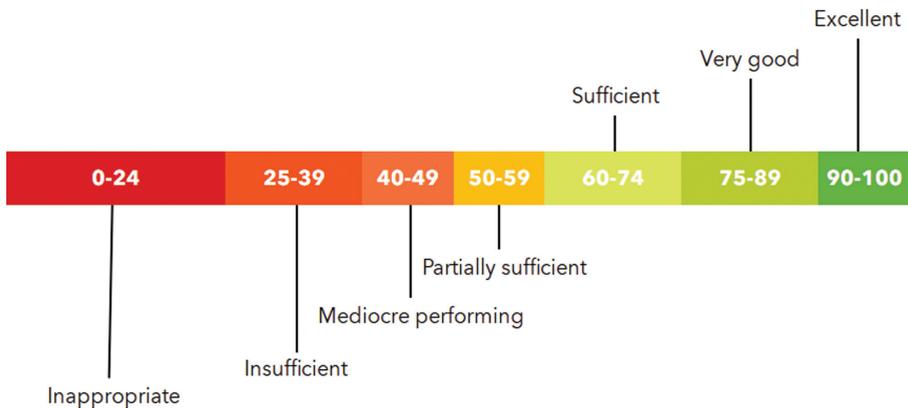


Fig. 2: Example of Mappings.
 Legend: Q = question, p = personnel.

| LIS | PSQ | p1 | p2 | p3 | p4 | p... |
|-----|-----|----|----|--|----|------|
| Qa | Q1 | | | personnel responsible for the provision of the service | | |
| Qb | Q2 | | | | | |
| Qc | Q3 | | | | | |

unbiased. The last three days of the period we distributed some printed copies (>50) to reach the final number of participants, which was 950.

To be able to compare the scores there should have been preceded two different mappings.

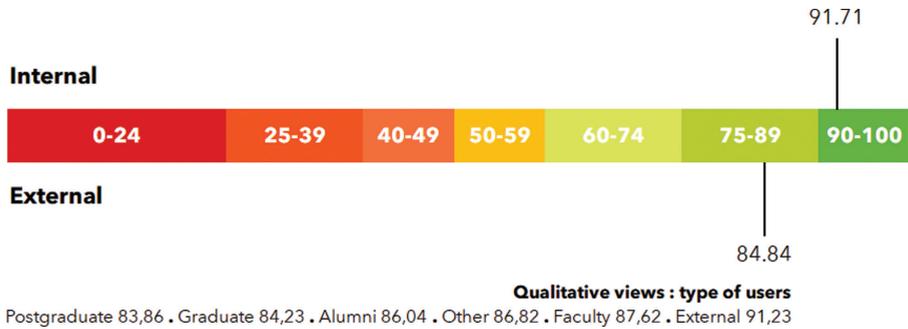
1. The first -horizontal- mapping was between common criteria in both instruments. This means that a criterion Qc in our instrument was conceptually mapped to the criterion Q2 of the PSQ instrument. To illustrate this, we give the example of question “Satisfaction rate from the visit and contact with the Library personnel” in our instrument, which we mapped to the PSQ item “Behavior towards citizens, as well as immediacy in serving their needs.” The mapping happened for three questions, one of which lead to the scores of five different services, e.g. satisfaction with circulation, support, instruction, etc.
2. The second -vertical- mapping was between criteria to resources. This means that we mapped the members of the personnel who are responsible for the performance of a service to the respective criterion.

Figure 2 illustrates the principles of mappings. It has to be noted that there could be no absolute mapping between all criteria of the PSQ and the user survey instruments, as some in the latter regard users’ satisfaction with facilities and collection that are not addressed in the former.

4. Findings

The use of the common scale provided us useful information on a number of criteria. The first comparison was the overall score of performance, which

Fig. 3: Patron Service Performance

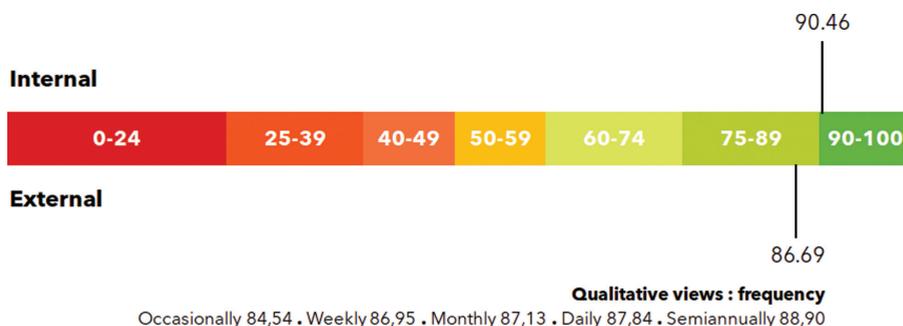


according to the first tier manager was 89.94% and to the user survey was 82.86%. This is a generic comparison that shows that the assessment scores, either internal, or external, correspond with each other. Although no statistical tests have been performed in order to validate this finding, the scores are indicating a consensus of high level of satisfaction with the overall performance.

For each one of these criteria we exploited the demographic data that we collected and we used cross-tabulation to see how these scores were distributed across several qualitative characteristics of our sample. Figure 3 presents the results for the criterion of patron satisfaction with service, which also refers to the mapping example between criteria. In this case we chose to see how the various types of library users evaluated our personnel’s performance in the area of service satisfaction and we found that external users and faculty members rate their experience with the service higher than the postgraduate and graduate users, who appear more demanding.

Figure 4 shows the comparison of rates for the satisfaction with the performance of the circulation service. As this item gives an example of the mapping between criteria and resources, the internal score was produced by the mean score of performance for all employees of the Circulation Unit, while the external reflected the satisfaction of our users by the performance of the service. In this example, we opted for an analysis of the frequency characteristics of our users. According to this, the users who visit our Library on a daily basis together with the ones that use the library once in six months are very satisfied by its services. These scores are approximate to the score

Fig. 4: Circulation Unit Performance



of the first tier manager, who believed that the Circulation Unit performed excellently. To further validate this we looked at the circulation figures, which were increased by 42.26% from 2015 to 2017.

5. Discussion

In all cases that we explored we found that this approach helped the library management to understand what is the level of performance from a wide perspective. It was the first time that we were able to interpret the results of a user survey with a scale that operates in a detailed and fixed context. As the rate for the evaluation is on the percentage scale, it was found easy to communicate to the public and understand it, while it was transparent and straightforward for the library administration to interpret the results. Furthermore, it was found that this approach provided a validation benchmark to the first tier manager for the comparison of his assessment practices.

We acknowledge that our approach has certain limitations. While the interpretation scheme is common, the rationale behind each score is different, as the viewpoint of each one who fills the instrument has a different understanding of the criterion. Thus, it was experienced that the mapping of the criteria between the instruments should be tread carefully, as the semantics of each question might not always be clear to the respondents. Moreover, all of these results should be interpreted in context, as there are certain limitations in resources that have to be taken into account and might not be known to the

public. Finally, as the data for the second tier manager are not accessible, this performance assessment is not fully complete. However, it is expected that in the forthcoming months, with the conclusion of institutional reforms, the Library will be able to administer itself the evaluation and to have full access to data from both tiers. As the evaluation of Public Sector has been normalized and is being conducted annually, this approach enables our intention to run biannual user surveys, following the same methodological way and having the same sampling target, so that gradually and in a light-weight mode, we can establish a comparison timeline for each criterion.

In a sense, in this study we have worked reversely to find a way to interpret what our users believe for our library's performance, by adopting the public sector scale and therefore aligning our various assessment data with the national context. However, in our view, national library evaluation initiatives should be informed on the applicability of international standards and toolkits. The fact that in many countries they are not used in a coordinated fashion, might be another potential reason why standards and toolkits, such as the ISO standards, have not been widely applied in the field.

6. Conclusions

In this study we approached the problem of heterogeneity with a solution that takes advantage of the existing percentage measurement scale for the assessment of Greek public sector employees to gather users' opinion on certain performance categories. This scale, as well as its interpretation module, was used for the harmonization of information that comes from varied assessment notions, tools and practices and secures the iterative and comparative nature of library performance measurement. With this harmonization, the administration of our Library is now able to see on a consistent and commonly applied scheme the various performance scores.

References

Arnau, R. C., Thompson, R. L., & Cook, C. (2001). Do different response formats change the latent structure of responses? An empirical investigation using taxometric analysis. *Educational and Psychological Measurement*, 61(1), 23–44. <https://doi.org/10.1177/00131640121971040>.

Association of Research Libraries. (n.d.) *LibQUAL+*[®]. Retrieved July 22, 2019, from <http://www.libqual.org>.

Association of Research Libraries. (n.d.) *MINES for Libraries*[®]. Retrieved July 22, 2019, from <http://www.minesforlibraries.org>.

Bertot, J. C. (2001). Library network statistics and performance measures: Approaches and issues. *LIBER Quarterly*, 11(3), 224–243. <https://doi.org/10.18352/lq.7643>.

Dole, W. V., Liebst, A., & Hurych, J. M. (2006). Using performance measurement for decision making in mid sized academic libraries. *Performance Measurement and Metrics*, 7(3), 173–184. <https://doi.org/10.1108/14678040610713138>.

Franklin, B., Plum, T., & Kyrillidou, M. (2009). Metrics in libraries and expectations for the evaluation of digital libraries. In G. Tsakonas & C. Papatheodorou (Eds.), *Evaluation of digital libraries: An insight to useful applications and methods* (pp. 17–39). Oxford: Chandos.

Greek Statistical Authority. (2019). Retrieved -in Greek- July 15, 2019, from <http://www.statistics.gr>.

ISO. (2013). *Information and documentation — International library statistics (ISO 2789)*. Geneva, Switzerland: ISO.

ISO. (2014a). *Information and documentation — Library performance indicators (ISO 11620)*. Geneva, Switzerland: ISO.

ISO. (2014b). *Information and documentation — Methods and procedures for assessing the impact of libraries (ISO 16439)*. ISO, Geneva, Switzerland: ISO.

Library & Information Center, University of Patras. (2019). *Statistics*. Retrieved -in Greek- July 15, 2019, from <https://library.upatras.gr/statistics>.

Ministry of Administrative Reform. (2016). *Human Resources Assessment — Legal framework*. Retrieved -in Greek- July 15, 2019, from http://www.apografi.gov.gr/evaluation/eval_thesmiko.html.

HEAL-Link. (2019). *Quality Assurance Unit of Academic Libraries (2019)*. Retrieved July 15, 2019, from <http://mopab.seab.gr/?q=en>.

Renard, P.-Y. (2007). ISO 2789 and ISO 11620: Short presentation of standards as reference documents in an assessment process. *LIBER Quarterly*, 17(3–4), n.p. <https://doi.org/10.18352/lq.7885>.

Richard, S. (1992). Library use of performance indicators. *Library Review*, 41(6), 22–36. <https://doi.org/10.1108/00242539210020541>.

Rudžionienė, J., & Dvorak, J. (2014). Public administration approach: What do libraries need for consistent information services evaluation, *Library Management*, 35(6/7), 495–507. <https://doi.org/10.1108/LM-02-2014-0019>.

Note

¹ An earlier version of this paper has been presented in the International Conference on Performance Measurement in Libraries, July, 23–25, 2019. Aberystwyth, Wales.