
Benchmarking user satisfaction in academic libraries – a case study

Sebastian Mundt

The author

Sebastian Mundt (37) is currently Head of Acquisitions at the University of the Federal Armed Forces in Hamburg, Germany. At the time of the survey, he worked as Assistant to the Librarian at Münster University and Regional Library, one of the libraries participating in the survey. He can be contacted at sebastian.mundt@unibw-hamburg.de.

Abstract

Customer satisfaction and service quality have so far been evaluated mostly from a local perspective although the quality element has been firmly established in academic library management for at least a decade. Critics of inter-institutional comparisons often object that different preconditions are not considered adequately. Examples from a joint user satisfaction survey conducted by 15 German university libraries in 2001 suggest that comparative data are a suitable means to identify cases of “best practice” and can effectively initialize processes of customer-focussed improvement. Furthermore, if compared with corresponding statistical data or performance measures, satisfaction ratings can reveal possible structural strengths and deficits relative to other libraries. On the other hand, follow-up telephone interviews with participating libraries showed that the survey results substantially challenged the institutions’ internal communication and public relations organisation, and underlined that even in a well-developed culture of assessment the need for professional mediation and coordination of comparative analyses may not be underestimated.

Objectives and organisation

All efforts put into satisfying library customers are based on the insight that services are not primarily chosen for their objective properties but first of all for their perceived fitness for purpose in the customers’ eyes. Satisfaction surveys are an established means to collect and gather these subjective judgements of single customers and convert them into a complex, objectified “snapshot”. In the process of planning customer-focussed improvements, however, it will still be difficult to assign reasonable priorities to the libraries’ services if the degree of satisfaction can only be rated on an isolated scale and not in relation to other institutions.

Having shaped a culture of assessment unique to the regional German library networks, all 15 North Rhine Westphalian university libraries decided to complement their regional statistical benchmarking system by conducting a joint user satisfaction survey across 34 branches in spring and summer of 2001. The aim was to gain an inter-organisational picture of user satisfaction which could be used by libraries to reveal and compare their individual profiles and improve the configuration of their services (Leibfried/McNair 1992, p.149). Since the early 1990s, North Rhine Westphalian university libraries have put a strong focus on researching and practicing performance measurement in libraries. Staff at Münster University and Regional Library and other institutions had made essential contributions to guidelines (Poll/te

Boekhorst 1996) and projects (e.g. EQLIPSE, EQUINOX) on performance measurement and statistics, and lately on cost management and controlling (Ceynowa 2000, Ceynowa/Coners 2003). In traditionally close association, colleagues of Dutch and North Rhine Westphalian university libraries have held a series of meetings to discuss principles and practice of performance benchmarking in their respective libraries (Poll 2000, Voorbij 2000). Yet, this prolific set of tools did not supply comparative data on user satisfaction.

The following institutions agreed to participate in the project and ensure methodological consistency:

The Hochschulbibliothekszentrum Nordrhein-Westfalen (HBZ), operating as the regional service centre for cooperative cataloguing, online interlending and digital library services, financed and supervised the project and served as “switchboard” between libraries.

infas - Institut für angewandte Sozialwissenschaft GmbH was contracted to develop a survey design, coordinate data collection and analyse the data. As a professional market research company with a strong nationwide reputation in the evaluation of customer satisfaction in public services, infas brought in extensive experience from previous customer satisfaction surveys and rankings in the public library sector (Windau 1997, Klug 2000). In cooperation with the Bertelsmann Foundation and the German Library Association, infas is also currently testing the tentatively called “BIX for academic libraries”, an annual nationwide ranking of academic libraries. Its set of indicators will contain a measure on user satisfaction.

Methodology

In accordance with their practice in previous benchmarking projects, the participating libraries set the following framework for comparing the outcomes of the survey:

- All data were collected under comparable conditions, i.e. by means of applying the same survey methodology. Despite the importance of a coordinated approach, however, it was agreed that any organisational conflicts had to be solved with respect to local requirements.
- infas staff conducted the input and analysis of survey data. This would ensure that the analysis was performed neutrally and professionally. infas hereby institutionalised the function of a fully respected neutral moderator between the libraries.
- Lessons can be learned even where different structures require diverging ways to suit the customers’ needs. Consequently, all results were presented to the participants without any pre-clustering or separation. It was left to the libraries’ discretion to consider and interpret, where applicable, different preconditions and environmental factors affecting the results.
- Each library was free to publish its own results; all comparative data, if used in public presentations, had to be anonymised.

Due to an overall lack of complete and consistent address data it was decided that the questionnaire should be distributed among library visitors and not as a postal, online or telephone survey among academic staff and students of the universities. It was accepted that this method did not match the criteria for random sampling as not every potential library user had the same non-zero probability of being part of the sample,

and that therefore the survey results could not be used to draw a general picture of all university populations in question.

Proportionate to the size of their respective populations to be served, infas distributed a total of 22,500 questionnaires to the libraries. In all libraries, 60% of the survey forms were displayed during the same three weeks in the summer semester, the remainder over a period of two weeks in the summer term break. Staff were asked to allocate them disproportionately to branches, i.e. displaying a minimum of 250 forms even to the smallest branch. Estimated from previous visitor surveys, and validated during the first days of the survey, the amount of non returns was estimated to reach about 40% of questionnaires on display so that a minimum of 150 returns per branch could be expected.

Staff members of each library were instructed by infas in a one-day training course on how to promote the survey among members of the university, to explain and disseminate the methodology to other staff involved and to organise the display and collection of questionnaires and the weekly return of responses to infas.

In order to raise awareness and inform about the background, objectives and schedule, the survey was actively advertised via the libraries' web sites and information display stands, as well as by newsletters, publicly displayed posters and other appropriate promotional materials in the weeks prior to the survey. This was deemed important to create awareness and willingness to participate especially among less frequent library visitors. To underline the seriousness of the project, the role and responsibilities of infas were highlighted on all announcements.

Library staff had to insure that questionnaires were on display at denoted places throughout the library's opening hours so that basically no person visiting the library should have to ask or search to take part in the survey. In some cases, predominantly occurring in large main libraries with gate counts of several thousand visitors per day, the response rate had to be adjusted by suspending the display of questionnaires on random days. In a few libraries with low returns from passively displayed questionnaires, staff were asked to actively approach users upon entering the library and hand out questionnaires.

"The ideal questionnaire is brief, easy to complete, unambiguous, and free of bias" (Baker/Lancaster 1991, p.371). With respect to these four success factors a previously developed questionnaire design was revised, coordinated, pre-tested and agreed between libraries. Given a restricted budget and time frame, and bearing in mind that only seven of the 15 participating libraries had conducted a user satisfaction survey before, the questionnaire intentionally did not adapt any scientific models of user satisfaction and service quality (e.g. Cook/Heath/Thompson 2001). Instead, it was set up in a conventional service- or product-focussed structure similar to, but more detailed than the SCONUL library satisfaction survey (West 2001). Users were asked to rate their satisfaction with 14 services, and to quantify the importance of each service on a five point Likert scale. The questionnaire has been reprinted and described in detail in Follmer/Guschker/Mundt (2002a). As the survey instrument was rather constructed and agreed than researched and modelled, specific elements of structure and contents were introduced to level local requirements to "standardize" the collection of data:

- The selection of questionnaire items mostly covered "standard" services offered by all participating libraries – manuscript and other special collections, for example, were not evaluated by means of these closed questions. Two open questions were intended to encourage comments on individual services and local settings and infrastructure.

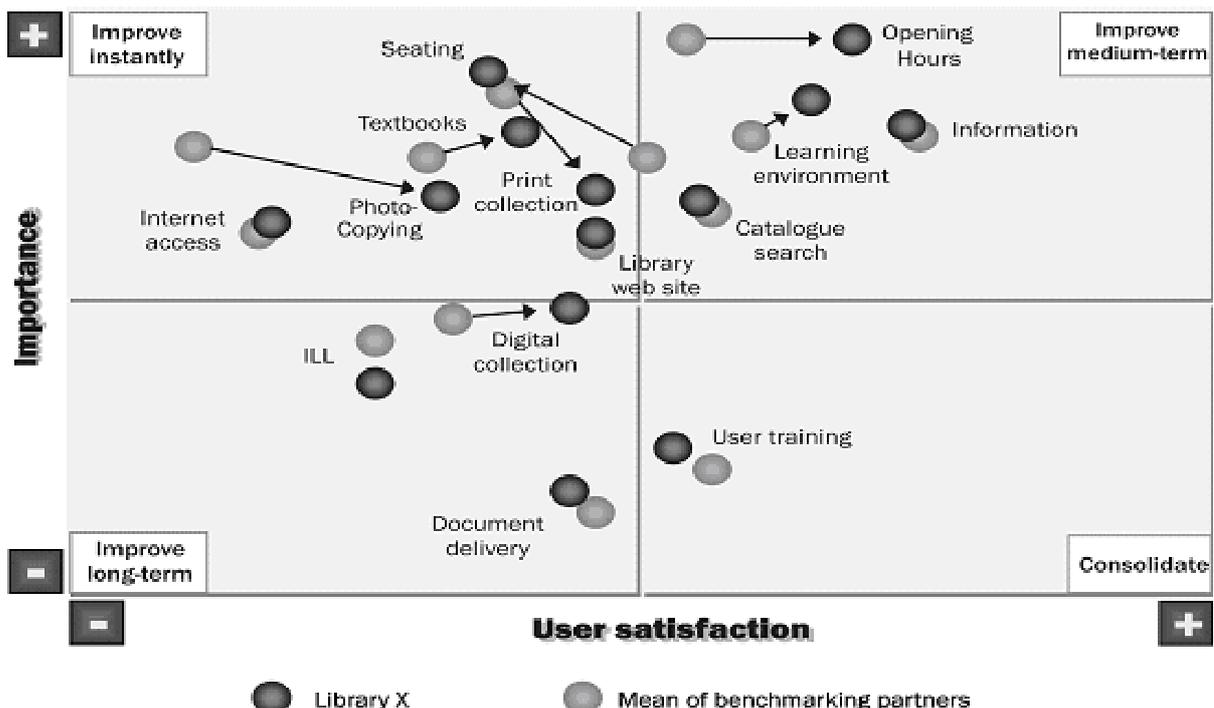
- Specific local designations (e.g. study fields, catalogues, abbreviations) were transformed into agreed standard terms and pre-tested as commonly understandable.
- Services that did not exist in all libraries (e.g. cafeterias) could be left out or deleted by respondents and were excluded from the analysis of the overall data.
- Each questionnaire contained an individualized grid chart displaying the respective library's standard opening hours. Respondents could tick boxes if and where they thought extensions were necessary.

Results and conclusions

By means of a factor analysis, the questionnaire items were grouped into mutually independent sets which were analysed for their plausibility. A regression analysis was then applied to determine the contribution of each factor to the overall satisfaction which had been asked at the end of the questionnaire. For the results of this analysis, the reader is referred to Follmer/Guschker/Mundt 2002b).

User satisfaction and importance of each service had been rated by survey respondents. In a first step, these two variables were displayed on parallel rating scales in a portfolio (see diagram 1). This representation served as an important tool to identify individual characteristics and „objectify“ the library's priorities for improvement. In this sense, priorities were not the services rated most important, but rather those which had been rated important and unsatisfactory – in the example shown in diagram 1 the Internet access, seating, photocopying facilities and the short loan collection must be regarded as the most critical gaps.

Diagram 1: Portfolio analysis of satisfaction and importance

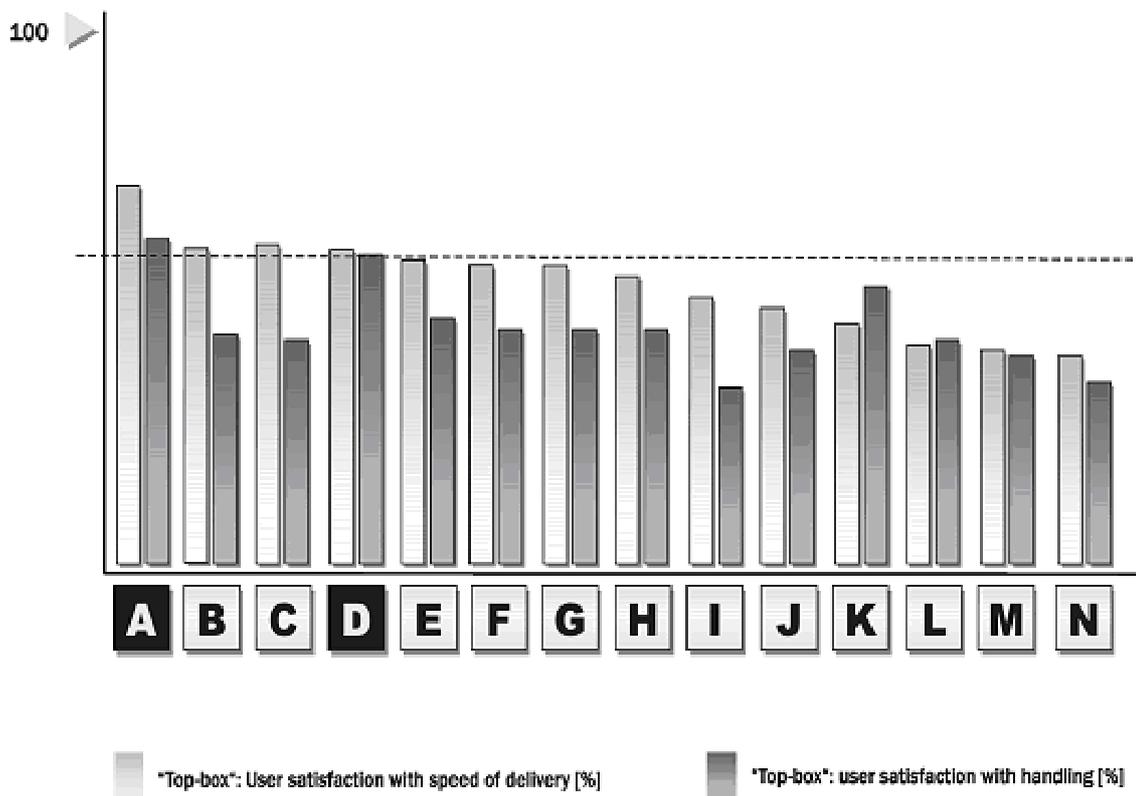


Furthermore, the representation illustrates specifics of one library's profile if contrasted with the average of possible benchmarking partners (represented in light grey). In this sense a principally "good" performance can be average, and a rather unsatisfactory grade can still be an example of better practice as seen in the example: The photocopying facilities are still being graded rather unsatisfactorily, but significantly better than in the benchmarking cluster. A close look at organisational details reveals that this library offers a high number of photocopiers and a flexible maintenance service to its users; these conditions have been negotiated independently from the university in a separate contract with the service company. The result seems to suggest a comparably good solution if this library cannot identify any better practice in other libraries. On the other hand, users of this library rate the seating facilities less satisfactorily than customers in other libraries do. As a study place, this library faces a particularly high demand as it serves staff and students from a university hospital.

The comparative search for best practice, however, needs a more detailed approach. For the analysis and understanding of processes behind best practice, other qualitative and quantitative information is often needed in addition to the primary information (user satisfaction) to interpret the results and identify possible success factors. This normally involves additional studies, surveys or observations; Town (2000) and Creaser (2003) have described such qualitative and quantitative analyses in detail. In the following examples, the regional online article delivery service JASON and the libraries' opening hours are used to demonstrate comparative satisfaction analyses.

Principally, JASON is being offered as a joint service; routing and supplier of a document will normally not be disclosed to the customer. In terms of satisfaction, the service gains its individuality by local marketing activities and an appropriate integration into a library's web site. The satisfaction with JASON has been divided into the factors "speed of delivery" and "usability" – the fee system was ignored as a factor as is used in all libraries. Diagram 2 displays both components; the so-called "top-box" value indicates what percentage of respondents were "satisfied" or "very satisfied" with the respective factor. In terms of delivery speed and usability, only two libraries (A and D) were able to satisfy more than half of their respondents in respect to both aspects. In addition, compared to other libraries, the JASON service has achieved the highest market penetration at library D.

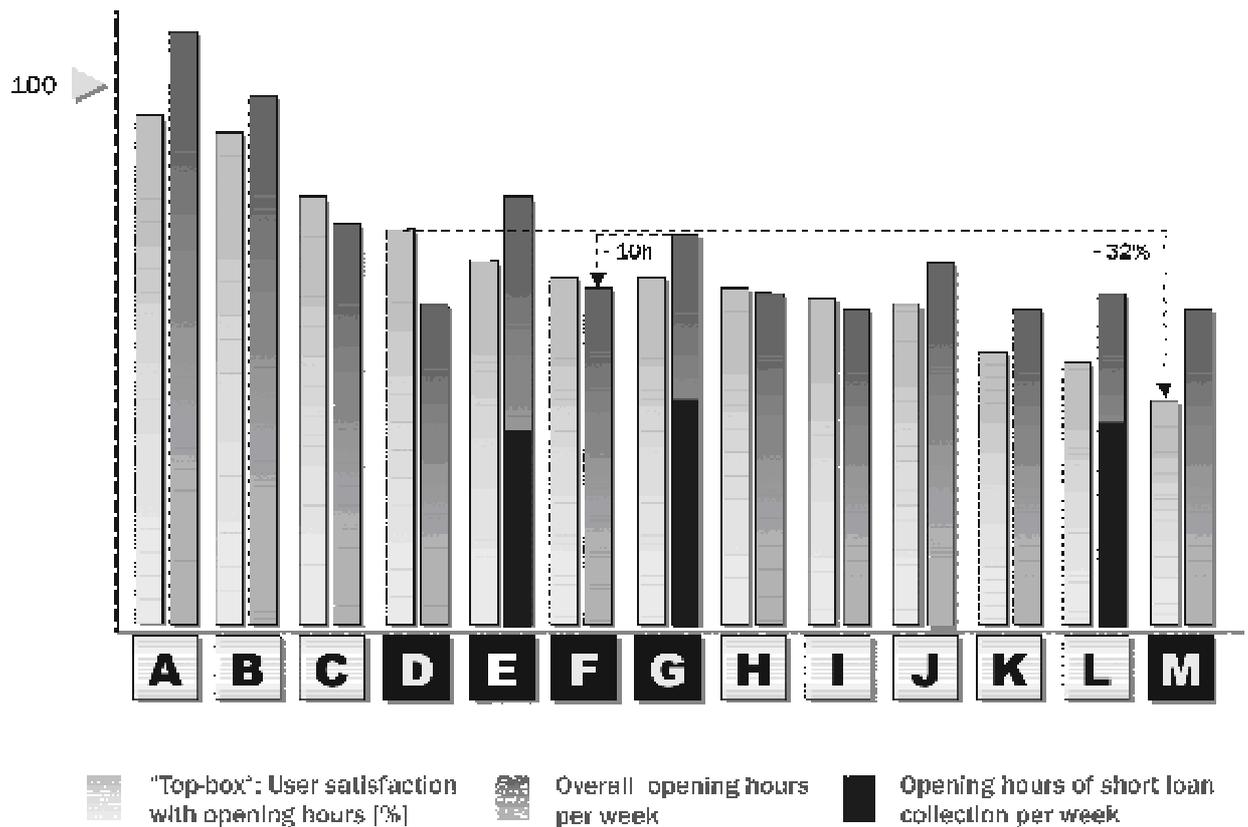
Diagram 2: „Best practice“ in regard to the document delivery service JASON



Part of both libraries' policy and success is an active marketing of electronic media and services which by far exceeds frequent user training lessons on JASON. Besides, essential differences between A and D and the other libraries could be found in the following characteristics: The high satisfaction with JASON in library A coincides with the recent successful implementation of an internal journal article delivery system to the academic staff members' desktops which has reconfirmed the library's competence for document delivery in the university. Library D has established a service point for document delivery and interlibrary lending in the heart of the main university building and effectively integrated the JASON order form into the progressive design of its web site.

Not all objects of comparison, however, necessarily lead to the identification of best practice. Diagram 3 juxtaposes the user satisfaction on the overall opening hours and the weekly opening hours of each main library. With about 100 hours of service per week, A and B satisfy 9 out of 10 of their respondents. From the graphical representation it seems obvious that the number of opening hours per week will generally have a strong effect on user satisfaction. About 70% of this influence can be explained statistically.

Diagram 3: Learning from complex environmental factors (opening hours)



Unsurprisingly, libraries offering restricted opening hours for certain services (e.g. short loan collection in E, G and L) faced significantly lower satisfaction ratings compared to those with similar overall opening hours. Besides this, long opening hours on weekday evenings generally had a stronger effect on user satisfaction than extended opening hours on weekends. Some outliers, however, show that the equation „longer opening hours = higher satisfaction“ can not be generally justified for all libraries (West 2001, p.10f.). D, for example, achieves 32% more satisfied respondents than M with the same total of opening hours per week, and G needs 10 hours more per week to satisfy its respondents to the same level as F.

To explain these differences, a number of infrastructural factors, e.g. the faculties represented, the location, the availability of public transport, the size of the institution and a number of other possible influences were examined. In the overall analysis, users from the fields of business and economics, law and medicine had been found to use the library more intensively as a workplace than those from other faculties (Follmer/Guschker/Mundt 2002b, p.23f.). J, L and M are the only libraries serving far more than one third of students from these fields; the opening hours of L and M, although representing an average level, satisfy less than half of their respondents.

An “environmental” indicator representing the difference between local opening hours and the mean value of other academic libraries in a zone of about 100 km radius was also found to significantly influence user satisfaction. It underlined a number of comments where users suggested proposals for improvement from their experience with other libraries. Customers of library G, for example, also use locations A and B whereas library F is located in the vicinity of L and M. Apparently their experience with service levels in other libraries will influence their expectations towards their

local institution. An overall systematic explanation of the respondents' ratings, however, could not be reconstructed from the data.

Follow-up

Four months after the results had been presented and put at the libraries' disposal, the participants were interviewed by telephone. The questions asked in this interview aimed at identifying typical procedures in the interpretation of data, the communication to stakeholder groups and the transfer of results into practice.

14 out of 15 participants used formal and personal channels to communicate the results to staff: nine libraries had informed their staff in employee meetings; two of them even chose to have the findings presented by infas project staff. Management in three large libraries disseminated the results in meetings with department heads. Two libraries planned an employee meeting at the time of the interview. In addition, 12 participating libraries made results available to staff via an Intranet or other internal media. In short, the majority of libraries had at this point introduced an important outcome and started a discourse about the results with members of staff.

In contrast, many libraries were still editing their results for public presentations; many of those who made slower progress were libraries without previous survey experience. An exhibition and/or web presentation of results had only been finished by one third of participants; five other libraries were in the planning or preparatory stages. Remarkably, one third of libraries abstained from presenting the results in the library although the survey had been conducted among visitors. Nine libraries had already reported the results to boards of the university. About half of the libraries kept a low profile with the external public, i.e. local press and media. One library even abstained from informing any stakeholders external to the library. The libraries' public relation policy, however, appeared to be unrelated to the survey results.

In terms of transferring the results into practice, the libraries provided a varied picture. All participants had derived their own strengths and weaknesses from the results. Only nine libraries, however, had set up a prioritised list of activities so far, and just a few cases of systematic benchmarking had been undertaken to study and scrutinize cases of best practice. At the time of the interview, 14 libraries had initiated activities – preferably those which would either quickly effect noticeable improvements or ones which involved long planning phases and/or complicated decision processes, e.g. building projects and other investments.

Discussion of findings

As a management instrument, benchmarking is not limited by the likeness or distinction of partners but merely by the adaptability of best practice to different structures. In order to participate, libraries only need to be "equivalent" in the sense that they want to profit from the results. In the absence of absolute standards, the opinion of its customers can be regarded as the most important quality feedback a library can receive. In exceptional cases, customers' expectations may be unrealistic or necessary financial, personal or infrastructural resources may not be available at the time or to the extent needed. Therefore, decisions in favour of the customer will have to be balanced against a library's resources and potentials (Ceynowa 2000). The need for learning and improvement is imposed by informed customers: The analysis of answers to open-ended questions indicated that users in this most densely populated German state with 15 universities are likely to assess facets of their

library's service on the basis of concrete experiences with other libraries. The satisfaction ratings with the libraries' opening hours underlined this indication.

In order to overcome the common saying of "comparing apples and oranges" and the inadequate public branding of "winners" and "losers", the libraries had decided to keep the results anonymous, to dispose of the dimension and objects of benchmarking themselves and to control the pace of progress. At the time of follow-up, however, a direct flow of information, i.e. communication with or visits to best practice libraries, was established only between a few benchmarking partners. Furthermore, the interviews showed that "progress followed organization": Libraries which had immediately installed project teams as part of a quality focussed and participating organizational culture were quicker and more successful to transfer results into action at the later stages. The "official" contract with infas ended with the presentation of results to all participants; in addition, a few libraries had invited the infas project managers to communicate and interpret specific results in local presentations at the libraries' cost. The presentation of results, however, only rings in the essential stages of a benchmarking process. In this sense, the follow-up interviews seem to indicate that especially the transfer of results into action needs more professional mediation and coordination to ensure a maximum benefit for all libraries.

The survey results were found to be consistent in the sense that those facets and services represented could be weighted as to their contribution to the overall satisfaction – not all relevant facets, however, were necessarily represented, and the questionnaire had not been tested and standardized as a survey instrument. Also, like web (Dillman/Bowker 2001) and other self-administered survey methods, the display of questionnaires to library visitors formally lacks the statistical requirements of *actively* selecting a sample of persons *representing* all members of the survey population. In this survey, the percentage of respondents visiting the library premises once per month or less often equalled 15%. It is likely that a further extension of the survey period would have contributed to a higher representation of infrequent visitors in the sample. Furthermore, the reason for infrequent or non-use of a main library or its branches might rather be a matter of (in-)convenience than of (dis-)satisfaction: In large universities, for example, faculty or other peripheral libraries offer important supplementary collections and services which were not represented in the design of this study. A special critique often applied to visitor surveys is that "virtual users" are underrepresented in the sample, and that results will therefore systematically underestimate the acceptance and use of electronic services. Results of a telephone survey among academic staff and students at Münster University in 2000 indicate that this type of user is so far mainly represented by a small, specific profile of academics and postgraduates specialized in Science, Technology or Medicine. Even if this type of users had amounted to 10% of the total population, the results of the survey described here would have changed only to a marginal extent. If libraries had been able to provide relevant demographic data and information on user behaviour, the results of the sample could have been weighted and "adapted" to the characteristics of the population. Due to privacy laws, however, no library was able to deliver the information requested by infas.

Nevertheless the ratings and comments from more than 14,000 respondents provided a differentiated picture which could be used effectively to discriminate targets for customer-focussed improvement. On the other hand, the effort to prepare, conduct and interpret a joint survey among 15 libraries must not be underestimated. Some

measures taken might even imply an initial stagnation or deterioration of service before users will actually benefit, e.g. when an air-conditioning system needs to be replaced or a new service contract for photocopiers has to be negotiated.

Finally, continuous improvement needs regularly updated results: In a rapidly developing hybrid environment, changes in user behaviour and the impact of improvement need to be re-assessed in regular intervals. Despite the initial one-off approach of the survey, this was consented to by the participating institutions. In 2002, the survey was revised and adapted for libraries of the universities of applied sciences in North Rhine Westphalia (Follmer/Guschker/Klitzke 2002). The introduction of the previously mentioned "BIX for academic libraries" is scheduled for 2004. In its set of indicators, user satisfaction is so far supposed to be measured separately by each participating library. Over time, the perceived political importance of rankings, and the demand for consistent data collection methods will probably induce libraries to coordinate the assessment of user satisfaction in academic libraries. It might even further the acceptance of an improved instrument incompatible to the results from this survey.

References

- Baker, S. ; Lancaster, F.W. (1991): *The measurement and evaluation of library services*, 2nd ed., Arlington, VA, Information Resources Press, ISBN 0878150617.
- Ceynowa, K. (2000): Managing academic information provision with the balanced scorecard: a project of the German Research Association, *Performance Measurement and Metrics*, 1 (3), 157-164.
- Ceynowa, K. ; Coners, A. (2003): *Cost management for university libraries*. München, London, K.G. Saur, ISBN 3598218346.
- Cook, C., Heath, F., Thompson, B. (2001): Users' hierarchical perspectives on library service quality: A LibQUAL+ study, *College and Research Libraries* 62 (3), 147-153.
- Creaser, C. (ed., 2003): *As others see us: benchmarking in practice*. Loughborough, LISU, ISBN 1901786536, URL: <http://www.lboro.ac.uk/departments/dis/lisu/op33.pdf> [2003-09-28].
- Dillman, D.A. ; Bowker, D.K. (2001): The web questionnaire challenge to survey methodologists, *Dimensions of Internet Science / U.-D. Reips and M. Bosnjak (Eds.) Lengerich, Pabst*, 159-178.
- Follmer, R.; Guschker, S.; Mundt, S. (2002a): Gemeinsame Benutzerbefragung der nordrhein-westfälischen Universitätsbibliotheken – methodisches Vorgehen und Erfahrungen [Joint user survey of university libraries in North Rhine Westphalia – methodology and experiences], *Bibliotheksdienst* 36 (1), 18-33, URL: http://bibliotheksdienst.zlb.de/2002/02_01_02.pdf [2003-09-28]
- Follmer, R.; Guschker, S.; Mundt, S. (2002b): „Alles, was man zum Lernen braucht...“ – Übergreifende Ergebnisse der Benutzerbefragung in den nordrhein-westfälischen Universitätsbibliotheken 2001 [“All you need for learning...” – overall results of the user survey in university libraries in North Rhine Westphalia], *ProLibris* 7 (1), 20-26.
- Follmer, R. ; Guschker, S. ; Klitzke, R. (2002): "Die Bibliothek bekommt ein Gesicht..." – Ergebnisse und Folgewirkungen der Benutzerbefragung in den

Fachhochschulbibliotheken NRW 2002 [“The library shows its face“ – results and consequences of the user survey in libraries of the universities of applied sciences NRW 2002], *ProLibris* 7 (4), 207-211.

Klug, P. (2000): BIX – The Library Index, or: Why less is often more, *Performance Measurement and Metrics*, 1 (2), 129-134; *BIX – Der Bibliotheksindex* [BIX – The Library Index], URL: <http://www.bix-bibliotheksindex.de> [2003-09-28]

Leibfried, K.H.J.; McNair, C.J. (1992): *Benchmarking – a tool for continuous improvement*, New York, HarperBusiness, ISBN 0887305482.

Poll, R. (2000): Three years of operating ratios for university libraries, *Performance Measurement and Metrics* 1 (1), 3-8.

Poll, R. ; te Boekhorst, P. (1996): *Measuring quality: International guidelines for performance measurement in academic libraries*. IFLA Publications 76. München, London: K.G. Saur, ISBN 3598218001.

Town, S. (Ed., 2000): *SCONUL benchmarking manual*, London, SCONUL, ISBN 0900210214.

Voorbij, H. (2000): Benchmarking in Dutch academic libraries, *Proceedings of the 3rd Northumbria international conference on performance measurement in libraries and information services*, Newcastle upon Tyne, Information North for the School of Information Studies, University of Northumbria at Newcastle, 253-258.

West, C. (2001): *Measuring user satisfaction: a practical guide for academic libraries*, London, SCONUL, ISBN 0900210257.

Windau, B. (1997): *Betriebsvergleich an Öffentlichen Bibliotheken. Bd. 2: Messergebnisse - Richtwerte – Handlungsempfehlungen* [Interlibrary comparison of public libraries Vol. 2: Results – reference values – recommendations], Gütersloh, Bertelsmann Stiftung, ISBN 3892042934.