

Setting standards for library and information service outcomes, and service quality

Rowena Cullen

Victoria University of Wellington, New Zealand

Whereas in the more traditional forms of library performance measurement focused on inputs, process and even outputs there is some acceptance of the value of setting base level standards, there has been little research in the field of impacts and outcomes, and no agreement on any form of standards. Research that can provide rigorous evidence of outcomes is needed for managers to make decisions that will maximise the impact of library and information services. Lessons may be learned from the health sector where research into the impact of information on clinical decision-making is being rigorously examined. The Evidence Based Librarianship movement proposes new standards for research that can be applied to outcomes research and also to the extensive work being done on service quality and satisfaction. The paper concludes that consensus on these methodological issues is necessary before consensus on standards for service quality could be promulgated.

Introduction

Library standards have in the past been primarily based on inputs, (financial, staffing resources, collection size) and to some extent on throughputs, or process/efficiency measures. Consensus on these, to the extent that it exists, (Moorman, 1996) has been based on comparisons of national and local statistics. Manuals of evaluation have recommended the use of outputs to evaluate library performance but output measures are rarely included in sets of standards. As the focus of research on the evaluation of library and information services shifts to impacts and outcomes of library services, and to satisfaction and service quality, the question arises of whether it is feasible, and indeed desirable to set standards in these areas as well.

The difference between the outputs and outcomes of information services, and user satisfaction and service quality is based on the nature of their relationships to information transactions. Outputs are the direct result of the interaction between inputs and process, that is, the transformation of inputs such as books, staff, and facilities into outputs such as loans, enquiries answered, reader education programmes attended. Outcomes are a result of these interactions, events and services. Outcomes are defined by ACRL as “ways in which library users are changed as a result of their contact with the libraries resources and programs” (ACRL College Libraries Standards Committee, 2000, p.175). Satisfaction and service quality also result from

interactions, events and services provided by the library. They are indicators of how the customer perceives the transaction and the service they have received, and are therefore a measure of the affective relationships that result from customer responses to these transactions. “For a library, service quality encompasses the interactive relationship between the library and the people who it is supposed to serve.” (Hernon and Altman, 1998). Indeed, Hernon and Altman (1996) go so far as to state that without the basics of service “no higher order outcomes (such as the enhancement of learning, the facilitation of research, or the impact on intellects) will occur.” The relationship between customer satisfaction and service quality is less well defined, service quality being both a component of customer satisfaction and a result of customer satisfaction with the service overall. The impact of service quality and satisfaction on outcomes also needs further exploration.

Outputs have traditionally been evaluated by means of reference accuracy and fill rates, title fill rates, document delivery rates etc. There are well-documented methodologies for evaluating these, and the debate focuses primarily on the unit of measurement and its validity (Van House et al., 1987). Outcomes have proved to be a more difficult area of evaluation, and there is no work to date on standards for outcomes, although there are areas in which it may be possible to provide guidelines and benchmarks for institutions interested in this form of evaluation. The area of service quality and satisfaction has been explored more extensively in recent years, and suggestions have been made concerning the possibility of using results of some service quality studies to set indicative standards of scores that libraries might include in their evaluation and planning procedures. Such proposals will be investigated in this exploration of the question of setting standards in the domains of outcomes and service quality in library and information services.

Output standards

Despite the work of people like Van House on methodology the establishment of both input and output standards for university and college libraries and public libraries has remained controversial. Standards for College Libraries promoted by the ACRL have shifted between quantitative standards sought by practitioners, and qualitative standards advised by “the ‘expert testimony’ of a small number of persons respected

within our field" (Parker, 1995). The current standards focus on inputs, outputs and ratios of inputs to outputs benchmarked against comparable institutions (ACRL, 2000). The ACRL has held back from any formulaic approach for university libraries. In both the 1979 standards, and the 1989 version the standards remain qualitative rather than quantitative. "A university library's collections shall be of sufficient size and scope to support the university's instructional needs and support the university's programs" (ACRL, 1979); "... collections should be extensive enough to support the academic programmes offered recognising that there are instances where reliance can be placed on access to other resources ... there should be provision for adequate funding for ... newly added disciplines..." (ACRL 1989). It is unlikely that the new revision of university library standards currently in preparation will be more prescriptive than this given the shift from collections to electronic services.

The Public Library Association has similarly moved away from prescriptive, input standards that began in 1933, culminating in the Minimum Standards for Public Library Systems (1966) and has refocused library evaluation on overall effectiveness (De Prosopo, 1973) and the standardisation of data collection for evaluation across inputs, processes and outputs (Altman, 1976). The current approach focuses on the use of planning procedures (Nelson, 2001) and techniques for measuring service quality and user satisfaction (Hernon and Whitman, 2001) as the main quality assurance procedures. "Individual planning and role setting thus replaced national standards as a way of measuring library services and effectiveness (Moorman, 1997). At state level, however, standards are still focused on resources, hours open and range of services offered because these are needed "when building the case for increased funding in budget discussions with governing bodies." (Moorman, 1997).

The Citizen's Charter in the UK is still primarily based on input standards, but national standards for public libraries out in draft format (United Kingdom, 2000) include a rather idiosyncratic mixture of inputs (materials expenditure, opening hours, range of services, staffing levels, workstations); services and processes; and outputs (primarily fill rates, and unobtrusive testing of the staff's and the collection's ability to answer queries, and user ratings of staff helpfulness). A maximum level (desired standard) is to be set for each where possible and a minimal level at which the Department of Culture, Media and Sport will intervene. Many of these have still to be defined. It seems likely that debate about final maximum and minimum values for a range of services, and for the more complex evaluation procedures will be protracted.

In the academic sector in the United Kingdom, by contrast, most of the focus is on standardisation of data for national databases of input/process statistics, benchmarking of outputs with comparable institutions,

and a number of EU projects focusing on new performance measures, including the EQUINOX project on performance measures for the networked electronic environment (EQUINOX, 2000). Most of these focus on input measures (percentage of population reached by electronic library services; number of library workstations available; percentage of information requests submitted electronically; total acquisitions expenditure on electronic library services), or process/efficiency measures (cost per session, cost per document, successful vs rejected sessions). Only one measure relates to the end results—user satisfaction is the last measure listed. It is possible that such a strong focus on inputs is a sign of an immature technology and that later measures will focus more on the actual success of users in retrieving needed information, and the outputs and outcomes of electronic library services.

Controversy therefore still surrounds the question of standards. Lancaster noted some time ago evidence that libraries make greater use of quantitative standards, in particular to obtain resources, than they do the qualitative standards and statements on policy and process that have largely replaced them (Baker and Lancaster, 1991). Despite the difficulties the profession has on many occasions in reaching agreement on such standards, he notes that there is value for the busy manager in having some concrete yardstick against which to measure achievement. It is for most a way of making sense of their world and their activities. Assessing user needs, planning for results, are activities that need to be put in some context, and given some meaning by external standards in some part of the process. Thomas Hennen in *American Libraries* put this question back on the agenda again last year criticising the profession for moving away from standards, and recommending a move not 'back' to basic standards, but forwards to:

- i) minimum standards for all public libraries in the US;
- ii) advisory, or target standards, possibly pegged to percentiles in performance that all should strive for, even if only a few achieve them;
- ii) benchmarks of excellence to be recognised as best practices (Hennen, 2000).

And yet uniform standards for inputs and outputs do not address the diversity of communities, and different roles of libraries in today's world. As Bob Usherwood comments on the new UK standards, they lack "an underlying philosophy" of library service, and the emphasis on measuring quantities rather than the imponderable matters of quality "makes it particularly difficult to rate services to people who are 'different' such as ethnic minorities and the socially excluded." (23 steps, 2000). The question of how to measure quality remains largely unanswered in the draft proposal.

Within the controversy a few common points emerge. We may well be on the way back to standards for input and process measures in libraries. Benchmarking is emerging as a useful form of quality assurance for the more complex outputs, while satisfaction scores and service quality measures may become routine ways of assessing the robustness of relationships between information services and their users, their responsiveness to stakeholders, and their competitiveness in the global digital environment.

Outcome standards

If the application of standards to what are now routine library measures is difficult and controversial, their application to outcomes, themselves still ill defined, is even more problematic. ACRL's definitions of outcomes, or 'ways in which library users are changed' include: academic performance improved by contact with library; library use leading to improved chances of a successful career; undergraduates using the library more likely to succeed in graduate school; students who use the library are more likely to lead full and satisfying lives. These have not been followed up by research, and contain too many other variables and 'confounding' factors to enable successful research design. If standards are to be set in the area of outcomes a great deal of work is needed firstly to establish units of data collection, and secondly to examine reasons behind varying results. Well-formulated and rigorously tested research questions and hypotheses are critical for robust results in the field of impact/outcome research. (Hernon and McClure, 1990, p.71-78).

Listing three of her own studies, in medical libraries, corporate libraries and the government sector that are among the few addressing outcomes as a means of demonstrating value to stakeholders, Marshall notes "the success of outcome studies depends largely on the researcher's ability to identify key activities that are directly related to the quality of a product or service and to create survey questions or focus group that explore the impact of information in these areas" (Marshall, 2000). The ACRL 'outcomes' need to be reformatted as researchable and valid questions in terms of Marshall's and Hernon and McClure's statements.

Despite the lack of rigorous research, there has been interest for some time in establishing outcomes, impacts and benefits of library services in the tertiary education sector. Powell (1992) summarises a number of earlier studies which attempt to investigate the question of impacts, largely associating library use and library instruction with user's perceptions of value, but concludes that research to date was inadequate to develop valid, reliable measures of the actual impact libraries are having on their users. Lindauer (1998) identifies important institutional outcomes to which academic libraries contribute, by looking at the standards promoted by the higher education accreditation

authorities in various states, and aligning them with library performance measures advocated in the LIS literature. She identifies some specific performance measures which provide evidence of this contribution, primarily output measures, and targets additional dimensions of performance that need to be taken into account in assessing the impact of libraries on institutional outcomes, for example, those favoured by key stakeholders in the institution as identified in Cullen and Calvert (1995), and the impact of a high quality campus networks (McClure and Lopata, 1996). These are summarised into five 'domains' in which the teaching-learning library is perceived to impact on institutional outcomes. While these studies provide valuable insights into the overall contribution of the library to educational outcomes none are directly focused on identifying actual measurable impacts on an individual or of individual transactions.

It is in the health sector, where most of the detailed work has been done on the impact of libraries and information on decision-making. The uniqueness of the clinical decision-making role and its independence from management structures, which confuse decision-making in the business sector, make this field a rich source of information about the impact of information provision. There have been a number of studies (Marshall, 1992; Klein, 1994; Burton, 1995) that demonstrate a change in a clinical decision as a result of timely and relevant information provided by an information service, and in some cases a cost benefit in reduced hospital bed stay.

Within the health sector the evidence based medicine /evidence based healthcare movement has resulted in increased scrutiny being given to 'interventions' based on bringing recent and rigorously researched information to the clinician at the point of care. Systematic reviews being sponsored by the Cochrane Effective Practice and Organisation of Care Group into the impact of paper - and computer-based information 'reminders' on clinical practice show some positive evidence relating to the benefits of information provision (Wyatt, 2000). However, even in the medical context Wyatt was able to find few examples of rigorous research into the impact of information systems to assist clinicians in decision-making. While the medical literature demands as its 'gold standard' the use of randomised double-blind controlled trials or at the very least a non-randomised controlled clinical trial for evidence of positive clinical outcomes of an 'intervention', researchers acknowledge the difficulty of developing similar methodologies which can demonstrate the impact of information 'interventions' in clinical processes, taking into account a range of variables impacting on clinical decision-making and eliminating bias in evaluating outcomes.

The value of health information impact research has been recognised in our own sector; rigorous methodologies from health research are being recommended

in the LIS sector to try to produce clearer evidence of impacts and outcomes. Hernon and McClure indicated in 1990 the significance of randomised trials to provide more reliable evidence of the outcomes of providing information to groups of library users (Hernon and McClure 1990, p. 71) although they acknowledge that randomisation of subjects to control and experimental groups is not always feasible in the library context, especially in applied rather than pure research contexts (Hernon and McClure 1990, p.83). Only in the area of information literacy research have there been studies that employ randomised or case-controlled trials of the impact of the intervention, although 'blinding' procedures in assessing outcomes are not generally employed.

More recently medical librarians in the United States and UK have formed an Evidence-Based Librarianship Implementation Committee whose first task is to define the "the most relevant and answerable research questions facing the practice of health sciences librarianship" and then to develop protocols for evaluating the validity and reliability of the evidence available from research, for developing practice guidelines and disseminating research results. Eldredge notes that this would require a shift in research methods from a theoretical approach based on history and philosophy, and methods derived from management and social sciences towards the more rigorously tested research designs of the health sciences, and in particular clinical medicine. The model is based on the wider range of methodologies encompassed by the Evidence-Based Health Care movement, rather than Evidence-Based Medicine with its disease-based model and insistence on RCT's as the 'gold standard', and argues for a "few intentional variations from the standard EBM approaches." (Eldredge, 2000, p. 290). Eldredge's proposed model for EBL, which seems likely to be accepted at the forthcoming EBL conference at Sheffield in September 2001, includes seven principles that include the following statements: "EBL supports the adoption of practice guidelines and standards developed by expert committee based on the best available evidence, but *not* as an endorsement of adhering to rigid protocol." "In the absence of compelling reasons to pursue another course EBL adheres to the hierarchy (or levels) of in Table 2 [see below] . . . for using the best-available evidence, lending priority to higher levels of evidence from the research." (Eldredge, 2000).

Like Evidence Based Medicine which suggests that 'expert opinion' and 'standard practice' are insufficient bases for clinical decision-making because they often lag far behind current best evidence, and that clinical pragmatism must be accompanied by evidence systematically gleaned from rigorously conducted research EBL allows for the application of pragmatic professional experience alongside a systematic review of the best research available as judged by a consensus of profes-

sional opinion on the appropriate criteria to apply to assess what 'the gold standard' is for LIS research.

The levels suggested by Eldredge, based on the ability of the research design to eliminate bias, form a hierarchy of value:

- Systematic reviews of multiple rigorous research studies,
- Systematic reviews of multiple but less rigorous studies (e.g. qualitative)
- Randomized controlled trials
- Controlled comparison studies
- Cohort studies
- Descriptive surveys
- Case studies
- Decision analysis
- Qualitative research

Library research, he notes, is focused on case studies, descriptive surveys and qualitative research. As methodologies in these areas improve, eliminating potential sources of bias and 'confounding variables' these methods may move to higher levels of evidence. Confounding variables in medicine include anything from aspects of a subject's lifestyle, diet, to other medications. In evaluating the outcomes of library and information services on specific experimental subjects these might include previous educational experience, educational programme engaged in, intellectual ability, learning style etc. Research designs that can control for such variables will be needed to establish the outcomes of library services as evidence.

Formulation of a researchable question, as Marshall notes, is the first hurdle. Just as it is extremely difficult in the health sector to ascertain changes in the health status of a population due to specific interventions without extensive longitudinal studies, but it is possible to demonstrate with a sufficiently rigorous research design the impact of a specific clinical intervention on a smaller population. We too need to lower our sights to specific outcomes that can start to build up a picture of overall impact. This can be done by testing the impact of specific interventions (e.g. research designs that I am currently working on will attempt to assess the impact on clinical decision making of training and desk-top access to Medline supported by no-cost document delivery, compared with access to Guidelines and protocols), but confounding variables such as clinical decision-making styles, and existing information-seeking behaviour of the clinicians involved are problematic.

Theoretically, testing a new information service involves the use of large samples and random allocation of subjects to the test and control groups, pre-testing to identify confounding variables, and using some form of a 'blinding' procedure in assessing impact, in order to avoid bias, but RCTs are not always ethically and operationally possible to apply in library settings (as Hernon and McClure noted (1990, p.80). Case/control group

comparison studies are more feasible and may represent a level of rigour that is an acceptable 'gold standard' within our profession. Assessing the outcomes of existing services may require the use of comparative studies between institutions serving similar constituencies (i.e. cases defined according to dependant and independent variables) but offering clearly identifiable variations in service in order to test for impact on some of the outcomes listed by Hernon and Altman (1996, p.2) most notably, the use of library materials by both faculty and students in their work, corrected for other variables such as information use related to the discipline, intellectual ability, study habits etc.

Nigel Ford and Andrew Booth suggest that as well as lacking the RCTs and other rigorous studies to provide the necessary evidence for EBL, our literature is not structured adequately (lacks systematic abstracts, and index terms that enable searchers to retrieve well conducted studies), and that we do not as yet have a body of systematic reviews such as are found in the Cochrane Library (which includes five databases of reviews and protocols). A focus on the highly structured research question (the outcome of a specific intervention in a controlled population) that informs medical research and retrieval of the results of research using similar questions is necessary to produce a body of systematic reviews that could be used for library decision-making. This again would necessitate the development of consensus on the criteria that should be applied to both quantitative and qualitative research that could be used as 'evidence' for informed problem solving and decision making.

The application of these ideas to the development of standards in relation to the outcomes of library and information services suggests that:

- i) standards that are to be applied across the sector for evaluation must be based on consensus concerning the value and relevance of rigorous empirical research;
- ii) there are few soundly conducted impact/outcome studies that there is no body of knowledge on which to base standards;
- iii) that consensus on the criteria for research would be a welcome first step in this direction;
- iv) that given the difficulty of designing randomised controlled trials to evaluate the impacts/outcomes of information services, a consensus of expert opinion might agree on comparative control vs case studies, as long as methodologies were able to demonstrate that the design would reduce bias and correct for confounding variables in assessing the outcomes.

Standards in service quality/satisfaction

Finally we turn to the question of service quality and satisfaction, measures of the strength of the relationship between the various component parts of an information service and its consumers, and the question of establishing standards in this area.

There are a number of ways of measuring both satisfaction and service quality. Two approaches which are currently dominating practice in the evaluation of library and information services in North America if not elsewhere are Peter Hernon's customer-focused manuals for the ALA on satisfaction and service quality (Hernon and Altman, 1996; Hernon and Altman, 1998; Hernon and Whitman 2001) and applications of the SERVQUAL instrument from the quality of service model of Zeithaml, Parasuraman and Berry (1990) in the ARL/Texas A&M trials of the LIBQUAL+ instrument. A number of other applications of the SERVQUAL instrument in libraries are referred to in Nitecki (1998) and Cullen (2001).

Hernon's approach to the question of standards in the field of satisfaction and service is based on his understanding of service quality as a multidimensional construct, which encompasses the content of the service (the materials, information or study space sought by the customer), the context (the experience of the user, interactions with staff, comfort of facilities etc.), and other dimensions such as the customer's expectations of service on that occasion, and any gap between those expectations and their perceptions of service, again on that occasion. However, as we noted above, satisfaction with individual encounters contributes to overall perceptions of service quality. It is also a result of the combined impact of these encounters that contribute to an overall level of satisfaction with the service in its entirety. Similarly, perceptions of service quality and satisfaction may be individual and collective leading to the overall reputation of the institution (Hernon and Altman, 1998: pp.8-9).

It is clear that these complex relationships between attributes of the service and attributes of users as individuals and members of multiple constituencies are on a quite different dimension from the impacts/outcomes examined above. However, these relationships form the basis of the organisation's success in the marketplace and should form the focus of evaluation in the context of library planning and decision making. Commitment to customer satisfaction and service quality involves ascertaining local customer needs and preferences, prioritising organisational activity in line with these and ascertaining how well these expectations are met. A variety of qualitative and quantitative research methods and analyses of institutional data that can be used to provide this information are outlined. A modification of the SERVQUAL instrument, explored in more detail in Hernon and Calvert (1996) and Nitecki and Hernon (2000) is included in the methods described

as an economical way of gathering data on expectations and perceptions of service delivery in the same instrument.

In this framework the role of standards belongs in the domain of inputs, processes, and some outputs, and need to be clearly communicated to the customer (Hernon and Altman (1996) quote an example from Wright State University Libraries Mission Statement). They are set for specific services by individual institutions in response to customer expectations (not necessarily unrelated to activities of other institutions and recommended professional standards where these exist). The focus is on the multiplicity of customer needs and expectations and the commitment to prioritising resources and focusing on whichever of these is deemed to be most crucial to the organisation's mission.

The quality service model of Zeithaml, Parasuraman and Berry (1990) provides the framework for the SERVQUAL instrument that attempts to capture some of this complexity with its measurement of customer expectations and response to 20 or more aspects of service delivery. The model itself is based on five 'gaps' that may lead to poor quality of service. These 'gaps' are analyse the discrepancy between: customer expectations and management's perceptions of these expectations (gap 1); management's perceptions of customer expectations and service quality specifications (gap 2); service quality specifications and actual service delivery (gap 3); actual service delivery and what is communicated to customers about it (gap 4); customers' expectations of service and perceived service delivery (gap 5) (Zeithaml, Parasuraman and Berry, 1990). Results from multiple applications of the original SERVQUAL instrument which was designed to analyse gap 1 showed that customer responses could be factorised around five 'dimensions' of service quality: tangibles, reliability, responsiveness, assurance and empathy. These dimensions represent the robustness of the institution's relationship with its customers on each aspect of service quality. Factorisation of the SERVQUAL scores derived from applications of the method in libraries has not consistently loaded on these five dimensions of service quality Nitecki (1998). Cook and Thompson (2000) suggest that three dimensions (responsiveness, assurance and empathy coalescing to form one dimension that measures 'affect') may account for the key factors in service quality in libraries, and indeed in other industries.

Cook, Heath and Thompson have continued to develop the SERVQUAL instrument in several studies at Texas A&M University, developing the survey instrument to include 19 additional questions intended to explore potential dimensions 'Access to collections' and 'Library as place'. (Cook, Heath and Thompson, 2001). In partnership with the Association of Research Libraries they have tested this instrument, now known as LIBQUAL+, across a range of institutions with the

intention of reaching consensus on the use of the instrument as a mechanism for setting 'normative' measures for meeting user's expectations. Analysis of the resulting data showed a very high level of item score covariance, and a robust single factor interpretation of the data. This raises the question of whether all these dimensions could be aggregated to an overall view of service quality and if user perceptions could be summarized in a single score reflecting one dimension only, characterized as "library staff understand and care about users' needs" (Thompson, Cook and Heath, 2000). The authors tentatively explore the question of institutional rankings based on such a score, acknowledging the need for further research before such a proposal is put to bodies such as ARL, commenting "It would be unwise to take the whole entire result set, place libraries in rank order and in isolation from institution specific context offer up that list as a definitive measure of service quality in academic libraries" (Thompson, Cook and Heath, 2000, p10). However, they do suggest that it is more likely that institutions with similar missions might find more value in studying each other's results (presumably across the full 41 items, and five new dimensions as well as the 'single score') in order to determine how well they were performing and where endeavours to lift service quality might best be applied.

To administrators and bodies responsible for accreditation this is a very attractive idea, not in essence too far removed from the use of satisfaction scores for staff knowledge and helpfulness in the UK Public library standards. Indeed many SERVQUAL studies have shown a strong correlation between overall satisfaction coefficients, and the existence of large and small gaps between customers' expectations over a range of services, and their perceptions of service delivery. However, there are several points to be considered here that need to concern us any intention to use this as some form of standard, or indeed to attempt to set standards in relation to SERVQUAL methodology. Is it possible to set valid and reliable standards for what are essentially measures of relationships?

SERVQUAL is essentially a diagnostic tool that attempts to determine the robustness of the relationship between an organisation and its customers along several dimensions. Damage to this relationship can be detected by the instrument and remedial action taken. It has value for longitudinal comparisons rather than inter-institutional comparisons, and any proposal to focus on a single measure would have to be treated in this context. Tentative suggestions concerning the 'meaning' of SERVQUAL scores that might indicate the extent of damage to the customer relationship along the five dimensions are reported in Cullen (2001). However, these relationship issues are unique to each individual institution, and examination of scores on each individual item is needed to ascertain where the problem lies, and what the remedy might be.

However, the quality service model of Zeithaml, Parasuraman and Berry (1990) explicitly allows for both organisational, and industry-wide specifications and standards for service quality and we should perhaps explore these in more detail before transforming the SERVQUAL instrument into something it was never intended to do. Investigation of potential remedies may involve investigations exploring the disparity between customer expectations, service quality specifications and actual service delivery — SERVQUAL gaps 2 and 3, using past research in areas such as reference service, focusing on known parameters of success in the reference interview (Cullen, 2001). An investigation of the impact on customer perceptions of service quality when setting individual standards for specifics of service delivery such as turn around times, document delivery times etc. (Hernon, 1996) or material fill rates as in the UK draft standards. There is more potential in using past research, and developing rigorous new research to establish industry specifications or standards for service quality that could be incorporated into a quality service model rather than attempting to transform the instrument itself, a diagnostic tool, into an industry standard.

Secondly there are continuing concerns about sampling methodologies and response rates in most SERVQUAL studies that are not completely allayed by the cogent arguments of Cook, Heath and Thompson in two other papers they distributed at the ARL Symposium “Measuring Service Quality” in October 2000 (Thompson, 2000; Cook, Heath and Thompson, 2000). Use of the web as a mechanism for delivering the SERVQUAL/LIBQUAL+ questionnaire highlights many of these issues concerning: the robustness of the instrument; the need to increase response rates and ensure the integrity of the responses; to ensure the research sample matches user population profiles in more aspects than role as faculty or student, and also reflects the ethnic diversity, socio-economic status, gender, edu-

cational background; and other issues (confounding variables) known to affect web based research, and which may also impact on library use and therefore perceptions of service quality. The Texas A&M team are to be congratulated on raising these methodological issues concerning their project so extensively but they have not yet been debated openly within the profession, and they must be. A far broader consensus on these questions is necessary before, as Eldredge notes, we can have confidence in the ‘evidence’ produced by this sector of research, and find some way of relating measures of service quality to measures of outcomes that will throw light on these problematic relationships.

In conclusion, it may be appropriate, and indeed timely to move back (or forwards as Hennen suggests) to standards for inputs, and possibly even process /efficiency measures, and some outputs. These can be useful management tools, and are valued by accrediting bodies. But they must be based on empirical research, as well as benchmarking with comparable institutions, and ‘expert opinion’ based on long experience. The maximum and minimum measures in the draft UK public library standards will need to be based on the same three factors. Our main endeavour here is to ensure that there is room for professional judgement, diversity and a sound empirical base for data collection. Other approaches to standards may be necessary in the more complex areas of relationships such as service quality and satisfaction, or outcomes where there is potential for confounding variables (e.g. cultural, and cognitive characteristics of users) or bias that can invalidate results. A first step would be to reach consensus on what kind of evidence we will accept in our profession as providing the benchmark we will use as our ‘gold standard’. Do we agree on the levels being set in the EBL group? Will we make a professional commitment to research at this level in order to enable to move more confidently to that ‘gold standard’ that fits our context and our needs?

References

- ACRL College Libraries Standards Committee (2000) "Standards for college libraries" in *College & Research Libraries News*, March 2000: 175-183
- ACRL University Library Standards Review Committee (1979) "Standards for university libraries" in *College & Research Libraries News*
- ACRL University Library Standards Review Committee (1989) "Standards for university libraries: evaluation of performance" in *College & Research Libraries News*, 50(8) 679-681
- Altman, Ellen (1976) *A data gathering and instructional manual for performance measures in public libraries*. Chicago, IL: Celadon Press
- Altman, Ellen; Beasley, Kenneth; and De Prosopo, Ernest R. (1973) *Performance measures for public libraries*. Chicago, IL: Public Library Association, American Library Association
- Altman, Ellen and Hernon, Peter (1996) *Service quality in academic libraries*. Norwood, NJ: Ablex
- Altman, Ellen and Hernon, Peter (1998) *Assessing service quality: satisfying the expectations of library customers*. Chicago, IL: American Library Association
- American Library Association and Public Libraries Association (1966) *Minimum standards for public library systems*. Chicago, IL: Public Library Association, American Library Association
- Baker, Sharon and Lancaster, F. Wilfred (1991) *The Measurement and evaluation of library services*. 2nd ed. Arlington, VA: Information Resources Press
- Berry, Leonard L.; Parasuraman, A.; and Zeithaml, Valerie A. (1990) *Delivering Service quality: balancing customer perceptions and expectations*. New York: Free Press
- Burton, Judith (1995) "The impact of medical libraries and literature on patient care in New Zealand" in *Bulletin of the Medical Library Association*, 83:425-30
- Calvert, Philip J. and Hernon, Peter (1996) "Methods for measuring service quality in university libraries in New Zealand" in *Journal of academic librarianship*, 22(5): 387-91
- Cook, Colleen; Heath, Fred and Thompson, Bruce. (2000) *How many dimensions does it take to measure users' perceptions of libraries?: a LIBQUAL+ study*. Paper tabled at the ARL Symposium on Measuring Service Quality, Washington, D.C.
- Cook, Colleen; Heath, Fred; and Thompson, Russell L. (2000) "A meta-analysis of response rates in web- or Internet-based surveys" in *Educational and psychological measurement*, 60 (6):821-836
- Cook, Colleen, Heath, Fred and Bruce Thompson (2001) "Users; hierarchical perspectives on library service quality: a LIBQUAL+ study" in *College & Research Libraries*, 62(2):147-53
- Cook, Colleen and Thompson, Bruce (2000) "Reliability and validity of SERVQUAL scores used to evaluate perceptions of library service quality" in *Journal of academic librarianship*, 26: 248
- Cullen, Rowena (1999) "Does performance measurement improve organisational effectiveness: a post-modern analysis" in *Performance measurement and metrics*. Sample issue: 9-30
- Cullen, Rowena (2001) "Perspectives on user satisfaction surveys" in *Library Trends*, 40(1):662-686
- Eldredge, Jonathon D (2000) "Evidence-based librarianship: an overview" in *Bulletin of the Medical Libraries Association*, 88(4):289-301
- EQUINOX. Library Performance Measurement and Quality Measurement System. Telematics for Libraries Programme Project LB-5634/EQUINOX 2447/10. <http://equinox.dcu.ie/reports/d2_5.html>
- Hennen, Thomas (2000) "Why we should establish a national system of standards" in *American Libraries*, March 2000: 43-45.
- Hernon, Peter and Whitman, John R. (2001) *Delivering satisfaction and service quality: a customer-based approach for libraries*. Chicago, IL: American Library Association
- Hernon, Peter and McClure, Charles R. (1990) *Evaluation and library decision-making*. Norwood, NJ: Ablex
- Klein, Michelle S. et al. (1994) "Effect of online literature searching on length of stay and patient care costs" *Academic Medicine*, 6: 489-494

- Lindauer, Bonnie Gratch (1998) "Defining and measuring the library's impact on campus wide outcomes" in *College & Research Libraries News*, 44: 546-569
- Lopata, Cynthia L. and McClure, Charles (1996) *Assessing the academic networked environment: strategies and options*. Washington, DC: Coalition for Networked Information
- Marshall, Joanne Gard (1992) "The impact of the hospital library on clinical decision making: the Rochester study" in *Bulletin of the Medical Library Association*, 80(2): 169-178
- Marshall, Joanne Gard (1993) *The impact of the special library on corporate decision-making*. Washington, DC: Special Libraries Association
- Marshall, Joanne Gard (2000) "Determining our worth, communicating our value" in *Library Journal*, 125(19):28-30
- Moorman, John (1997) "Standards for public libraries: a study in quantitative measures of library performance as found in state public library documents" in *Public Libraries* 36(1) 32-39
- Nelson, Sandra (2001) *The new planning for results: a streamlined approach*. Sandra Nelson for the Public Library Association. Chicago, IL: American Library Association
- Nitecki, Danuta A. (1998) "Assessment of service quality in academic libraries: focus on the applicability of the SERVQUAL" in *Proceedings of the second Northumbria international conference on performance measurement in library and information services*. Newcastle Upon Tyne: Department of Information and Library Management, University of Northumbria at Newcastle
- Nitecki, Danuta A. and Peter Hernon (2000) "Measuring service quality at Yale University's Libraries" in *Journal of Academic Librarianship*, 26(4): 259-273
- Parker, Diane C (1995) "Standards for College Libraries: foundations" *College & Research Libraries News*, May 1995:330-331,337
- Thompson, Bruce (2000) *Representativeness versus response rate: it ain't the response rate*. Paper tabled at the ARL Symposium on Measuring Service Quality, Washington, DC, October 2000
- 23 steps to pleasing the DCMS (2000) *Library Association Record* 102(6), June 2000: 303
- United Kingdom Department for Culture, Media and Sport. Libraries, Information and Archives Division (2000) *Comprehensive, efficient and modern public libraries—Standards and assessment* <www.culture.govt.uk/heritage/index.html>
- Van House, Nancy et al. (1987) *Output Measures for Public Libraries*. 2nd ed. Chicago, IL: Public Library Association, American Library Association
- Wyatt, Jeremy (2000) *The effects of manual paper reminders on professional practice and health care outcomes: a systematic review of 22 randomised controlled trials including 41705 patients*. Revised final project for NHS R & D project no. IMP 15-11, 14/2/2000