Cheap, useful and fairly valid performance indicators for tertiary libraries

by Isabella Trahn

Paper delivered at the the 20th Annual Conference of the Australian and New Zealand Theological Library Association 8 July 2005, Veech Library, Strathfield

Introduction

Some of you may think that what happens in performance measurement at a very large university library such as UNSW is not so relevant to you because of differences of size and complexity. In fact, much of what the largest libraries do in this area does not have a particularly high technological component. The overall pressures faced by both large and small academic libraries are essentially the same. All library staff have "day jobs", so when it comes to getting the information to manage a tertiary library what is needed is to get and to use relevant information as quickly and painlessly as possible.

This paper looks, in the main, at quantitative measurement.

Cheap, useful and fairly valid definitions

Why this title? Some of you may recognise that the adjectives in the title of this paper can be attributed to Swinburne University's Derek Whitehead. Derek wrote a commonsense paper on university library statistics in 2003 using those three practical qualifiers. The PowerPoint slides of Derek's paper can still be accessed through the CAUL Statistics website (www.caul.edu.au)[i] which is a useful website recommended to anyone interested in performance assessment both for its content and for useful links.

* Cheap
* Useful
* Fairly Valid

To these three I would add three more necessary adjectives. Performance measurement in its practical application within our time-impoverished work lives needs to achieve worthwhile purposes or it is not worth doing. Our working hours are too short and too crowded to do otherwise. The three additional descriptive qualifiers I would add are:

* Contextual
* Selective
* Common sense

Cheap

Cheapness, of course is a very subjective quality. The eye-of-the-beholder factor is influential and cheapness is not an absolute. One library's cheap is another's prohibitively expensive. To undertake a full Rodski[ii] survey is expensive, and to do it regularly adds up to a tidy sum. On the other hand, allowing a large catalogue of user woes to remain not properly identified or addressed may result in a vastly more expensive outcome for the library in the medium term. When user support is vital and university resource allocation starts to turn Darwinian, a little investment in data as ammunition for the library cause can reap dividends. Thus cheapness and usefulness are linked. The more useful a set of indicators, the more reasonable a resource outlay to provide these may appear to be. Value for money may be a better description than cheapness. Positive cost benefits in terms of the use of your time makes the relevant performance measure more attractive.

Isabella Trahn, B.A. (UQ), Dip. Lib.and M. Lib. (UNSW) is the Quality and Planning Librarian, University of New South Wales Library. For the past decade she has had in her portfolio, responsibility for, and an interest in benchmarking and quality issues, statistical reporting and surveys for UNSW Library.

Of course, cheap may be best in the first place and most fit for the purpose. Within a specific service excellent results can be achieved, for example with something produced and implemented at absolutely minimal cost such as a simple, well designed page for feedback together with an effective, low key, timely way to deliver and retrieve the information. Two recent examples, one implemented, one not, for other reasons, of potentially cheap qualitative data gathering of this type in our Library illustrate the role of simplicity. The first proposal was a small sheet, distributed and collected by security staff as part of their supervision duties in a pre-examination study hall, extending the hours of opening period. Ultra cheap, not costing extra staff time, these sheets have the possibility to provide very valuable feedback, basic statistics and a range of grateful user comments appropriate both for the improvement of the study hall mode and for providing strong user endorsement for the relatively new policy. Result – cost marginal, benefit positive. The second was a simple form collecting brief reflective comments by staff on out of hours shifts at reference desks to describe their experience of the volume and type of user during out of hours opening. This built up, in conjunction with simple quantitative data, a distinct picture of user behaviour, despite individual staff differences in reporting business.

As with a budget area such as staff development, if you have large resources available, a whole range of ambitious activities can be provided, but if you simply don't have the budget, it is still surprising the range of opportunities which can be provided using imagination, ingenuity and the resources around you, particularly if your Library is set within the university context, on a shoestring budget. Many years ago I presented a staff development program for 200 staff, for six months with a zero budget, apart from salary, using university expertise, then available without direct charge.

Usefulness

Commonly suggested uses for performance/ indicators include:

- Aligning services with institutional requirements

- Monitoring trends
- Seeing if you are at the top or bottom of the peer totem pole which may indicate self-congratulatory celebrations or, alternatively, desperate improvement measures are in order
- Basing changes in policies and procedures on real data
- Identifying specific areas for priority improvement
- Making claims for additional resources
- Justifying resource expenditure
- Exploring the level of user satisfaction with services
- Investigating the relative importance users place on services
- Investigating how services and resources are being used

There are many more worthy reasons and I'm sure you can add many other reasons to these. We will return to these practical outcomes again later.

Steve Hiller from the University of Washington Libraries, who was in Australia recently for the Victorian-based training provider CAVAL, considers the following to be useful questions to ask. "Do we have the data? Do we need the data?" These two questions nicely encapsulate situations in individual libraries where there are always gaps in data collection just when and where you really need them and yet where the volume of existing data collection would indicate that there may be some unnecessary work going on – just because it always has. Steve Hiller and Jim Self from the University of Virginia Libraries are currently American Research Libraries Visiting Program Officers undertaking a national program to raise the level and quality of Library performance assessment in US research libraries. The ARL website another recommended source of data for those interested in performance measurement.[iii]

Perhaps it is no longer useful to collect data you have been collecting forever. Examples of "traditional" local counts which one of our departments had been meticulously keeping for 30 years but which are not longer collected are "mail items opened" and "journal issues entered". These counts show the busyness (or perhaps lack of busyness) of a small part of our Library, but how many mail items opened is

good performance? And what does that have to do with supporting research and learning and teaching? It is the business of the Library which is important, not the busyness.

In these days of fewer staff and staff over-commitment for almost everyone in the tertiary sector, the temptation to spend valuable time documenting busyness still persists and is ever more tempting to try to justify our current overworked state. Our Library departments tend to do this monthly in their reports which slide into compilations of busyness with little comment on outcomes or issues of importance. University administrations tend not to be very sympathetic when what we are saying is Look see how much activity we have! Look at our dedicated staff! 'Beyond busyness' is a more useful area to focus upon.

Fair validity

If all Library managers were completely honest they would admit that many of the measures and statistics that they produce are often imperfectly designed and collected or based on historical figures of very dubious validity. One wonders how many of the "how big" measures are even close to reality. Some of these have been maintained for a very long time. What do you do? Take last year's figure, add items added, subtract items removed – and there is the result. Your addition and subtraction may be fairly accurate but what about the base figure lovingly handed down from fifty or a hundred years ago? How valid are totals combining films and books, microfiche and tapes, sheets of maps and DVDs? Is a 900 page medical dictionary equally valid to a microfiche with three images on it?

Contextual

As much of what is discussed here indicates, figures by themselves are fairly meaningless. Only where the context denotes some form of comparability will some figures begin to make sense and the possibility of any type of comparison between institutions become even remotely possible. Looking at roughly like institutions tends to be done by broad groupings and categories. This type of approach is demonstrated by loose Australasian groups of tertiary libraries such as:

Group of 8
Universitas 21
ATEN libraries
The gum leaf universities
NULG
ANZTL.

Our Australian Defence Force Academy campus is a good example of a library sensibly looking towards its natural peers to continue its performance measurement. ADFA looks at, and is developing relations with, smaller military libraries around the world rather than looking to the UNSW main campus library for their peers. Innovations, ratings, favourable comments are all noted carefully. Most university libraries have a very small number of "rival" institutions with which surreptitious or overt comparisons are always made. In the UK there is an 'ancient' vs. 'recent' divide and in the US overall size enables banding of broadly like institutions. Whatever the common criteria, look across the world, your peer group may well be global. Don't just look around the corner. A good grouping is also not so large that active communication between peers is difficult.

This is all part of the process of looking for context and perspective. The Society of College, National and University Libraries (SCONUL), the organisation of UK university and national libraries, suggests that all statistical statements be accompanied by contextual information and an example of this approach is demonstrated here by the Council of Australian University Libraries (CAUL) which finishes the annual statistical survey with the institutional data which it takes directly from the Department of Education, Science and Training (DEST) in order to minimise reported variations which might distort the other statistical items. The suggestion is that nothing should be discussed without contextual data. Complex institutions are so much themselves that it is unfair to compare otherwise.

A word about selectivity

One interesting development in the UK some years ago was the emergence of projects to identify not more than one dozen key performance indicators which could be used to solve the issue of information overload for university managements. The principle behind this project is a good one. Find what

works. Make sure it is clear. Make sure it is appropriate, and enjoy the focussed effort rather than trying to maintain too broad an approach.

Common sense

I don't think there is any need to expand on common sense for this audience.

Performance indicators and measures

So much for adjectives. As for the noun, there are many definitions of performance measures. For the practising manager, context in performance measurement is also the key. My personal preference for a definition of performance measurement and indicators is one which I quoted in Guidelines for the Application of Best Practice in Australian University Libraries; intranational and international benchmarks (Canberra, DETYA, 2000)[iv] which I co-authored with Anne Wilson and Leeane Pitman. The definition is by Pieter Te Boekhurst who has worked closely with Roswitha Poll at Munster University Libraries. Both Roswitha and Pieter have published extensively in the area of statistics and performance measurement. Te Boekhurst says:

"Performance measurement is comparing

WHAT A LIBRARY IS DOING
(PERFORMANCE)
With
WHAT IT IS MEANT TO DO (MISSION)
And
WHAT IT WANTS TO ACHIEVE (GOALS)

The extent to which goals are reached can be determined by using performance indicators."[v]

I mention The Guidelines publication because it is a useful summary of the state of performance measurement in Australia and beyond at the time of publication over 5 years ago. One of the themes of the chapter on performance measurement and a concept which I adhere to in day to day management is indicated by the definition above. That is that indicators of any sort only have meaning in the organisational context. The more I have had to do with benchmarking over the years the more I can see measures tempered by the

nature of individual institutions and individual libraries. Apparently striking differences or apparently striking similarities can both be incredibly deceptive. Like individuals in a team, organisations have their own DNA, if you will and many valid reasons for particular outcomes.

The judicious use of measures or indicators as shown by the definition above is not only a mechanism for day-today management and improvement, it is also a way to explore the impact of services, to test out wider options, to justify resource allocation and potentially to benchmark with others. As such it is an important part of strategic planning. Out of context, an indicator indicates nothing useful. This relates well to the old quality adage: FITNESS FOR PURPOSE: doing what is right for US. I was reminded of this yesterday at the Australian Universities Quality Agency conference here in Sydney, where the notion that AUQA audits do not critique organisational goals, but look at how fit the organization is for its stated purpose. Are they doing what they say they do?

Performance and statistics

Looking at first level sources of data for performance measures brings us to the fascinating area of management statistics. Counting, often and widely, is the oldest and most widely practised method to try to describe size and busyness of libraries, so no discussion of performance measures can omit statistics and their derivatives.

Mine is bigger than yours

Bland figures indicating institutional and/or library size (institutional population, budget size, staffing composition, floor space, number of seats, numbers of computers); information resource dimensions (holdings of physical and electronic resources) and/or service use and patron perceptions (circulation, accesses and downloads for electronic resources, service ratings in client surveys) build a picture which may be used in a managerial sense for many purposes but this picture only realises its true potential as a source of sensible library performance indicators when it can be shown that, as Te Boekhurst says, the picture of PERFORMANCE is firmly set within the context

of MISSION and GOALS.

More than the numbers

What does having a budget of $22 million for your Library tell you? Does it mean that, if another Library has a budget of $30 million and has a similar institutional population they are remarkably better off? It may be that the second institution has a whole range of disparate campuses and must provide services across time and distance and in a profile of uneconomic small clusters of services. It may be that "total budget" figures may, or may not include a range of discretionary funds. There may be, and usually are, a whole range of reasons why what seems apparent, is not necessarily so. The apparently most straightforward of figures cannot tell the whole, or even the necessarily true, picture.

The threatening nature and power of numbers can be illustrated by a story. Some years ago the four Australasian members of the international university grouping which goes under the title "Universitas 21" spent some time putting together common Library profiles. The profiles were long and detailed and the effort expended not inconsiderable. The University Libraries concerned were Queensland, UNSW, Melbourne and Auckland. It was fascinating to watch the then university librarians looking at the "comparative" figures diligently compiled by their staff for the purpose. Those four university librarians could see so many "yes but…" and "this doesn't include" and "but you do x…." interpretations of the most basic reported aspects that there was the potential for whole segments of meetings to be contained in discussing caveats. After a few short years, maintaining the full profiles lapsed, although individual libraries did continue using the format as they found it internally useful. I suspected particularly that figures which reflected favourably tended to be believed and those which did not, tended to be seen to contain flaws.

Counting busyness

It is natural that Library staff, who serve users with such dedication over many years, for mainly intrinsic rather than financial rewards, wish to prove that they are not idle, that their efforts have encompassed all this work.

Measures of busyness and activity may be used for self-congratulatory purposes to indicate how much "more has been done with less", but busyness is a very overrated measure. Busyness is better expressed in terms of the allocation of staff time spent. Most universities do not want to know how busy Library staff have been. They want to know whether the mission of the library is being fulfilled, how much "bang for their buck" they are getting. The real question is whether the proportions of staff time dedicated to particular activities moves the organisation forward. Is the effort going where it should be? At another level, so does AUQA. Are we doing what we say we do and how do we demonstrate this?

Classic busyness: What is "Number of reference inquiries"

The "how big" material figures can deceive but the "how many" service figures can be a whole other dimension of things not being as they seem. What is an "inquiry" or "reference" statistic and can numbers of (reference) queries tell you anything? CAUL sensibly did not collect this figure as a compulsory field for many years. Each institution, and even units and individual team members within each institution have an idiosyncratic accretion of definitions of "reference".

For example, in our Library, for many years, statistics were kept of "inquiries at reference desks." Then, aware that these figures covered a multitude of sins, the collection of these was further differentiated into "reference", "directional" and "technical". Even within these there are debates about what is/should be classified under "reference". We know there are discrepancies in what is recorded where, for service desks just one floor apart. What we also know is that the idiosyncrancies within each area are fairly consistent over time because of local unit traditions. Individual desk trends over time are thus of value, but even they have to be related to events such as changing desk hours and do not tell a simple story. Some staff maintain inquiries may be fewer, but they are more complex. Other measures are needed to illuminate whether this perception is fact. More of this below. Questions need to be asked. What else is going on? Inquiries are declining, but what about information literacy figures – are they

going gangbusters? It's useful to have more to give university administrators than figures in decline. Unfortunately, unless your Library has its e-reporting in place, that is exactly the direction for the bulk of current traditional statistics.

The usefulness of numbers

So, should CAUL stop collecting the AARL statistics and the ARL in the US discontinue their 60 year tradition? No, they shouldn't. What both those bodies have done is to create large data banks and some tools which now enable useful (our second adjective) customised analysis at point of need, for a specific purpose. CAUL have followed the lead of ARL and enabled interrogation of their statistics database by individuals, when needed. This is a practical exercise in efficient individual control and provides hitherto cumbersome opportunities for the exploitation of the data. Customised use makes the collection and collation of the data more worthwhile. The new developments on the CAUL Statistics site kept for so long have enabled DIY trend analysts to customise the bank of statistical data carefully and comprehensively .

The real value in any statistical figures lies in trend patterns, especially over time, but also across a sector. Trend analysis has become easier to maintain and provides far more visual impact with each improved version of Excel. Within the Library, or particularly in dealings with university management committees with short attention spans and a lifetime of assumptions about academic libraries, the power of the image is exceptional. Eyes which glaze over when faced with banks of tables or which become absorbed in an academic analysis numerical minutiae, lighten up at a bank of rising or falling bars or galloping coloured segments of pie charts.

The power of the image encapsulates dull statements in exciting ways and shouldn't be underestimated.

The usefulness of indicators: simple combination indicators/ratios

Over the years CAUL has responded to the realisation of the limitations of "how big" and "how busy". CAUL has introduced next stage indicators of the dollar per population member variety which combine two streams of data to give simple measures which can be ranked.

Look at the ratios on the CAUL Statistics site; they are much more dramatic in terms of comparisons and the ability to grasp the wide range of resource richness and poverty between Australian university libraries. These indicators have a strong corrective property. Number of seats means nothing. Number of seats per student starts to mean something. Dollar expenditure on journals may not mean a whole lot, except a sharp intake of breath by a university administrator, but dollar spent on serials per academic staff member/researcher may mean a little more.

These measures generally bring the user into some perspective in terms of resource allocation averaged over all such individuals. It is easy to picture one user with a large pile of dollars, many seating options, groups of eager staff available to assist them, and then another user, with a tiny pile of dollars, half a chair and librarians no where to be seen. Tables of large numbers are harder to embrace. Simple ratios are a favourite category of mine. They also tend to have management impact.

Another approach using simple combinations includes one recommended for SCONUL librarians a few years ago. This was to combine measures for reference desks and information literacy and to express this in terms of overall staff hours dedicated to these activities rather than attempting to count queries and information literacy sessions. Since many librarians see so-called reference as a specialised type of information literacy activity, this makes sense in that context. Staff hours is a measure underutilised but one administrators understand. Your Library spends $12 million per year on staff. They are paid x at this level and y at that. These sorts of figures mean more than statistics such as 250,000 queries answered per year. Staff activity measures of this sort begin to educate management on the hidden costs of, for example, information literacy. It takes how much staff time to prepare for eight hours of in-house course instruction?

CAUL Performance Indicators

Most of you would be familiar with the three

longstanding CAUL indicators:

1. Client Satisfaction
2. Document Delivery
3. Materials Availability

These were developed by CAUL, adapting models from elsewhere in the early 1990's. By 2000 at least 25 of Australia's universities had used or were using some or all of these irregularly or regularly, in their search for management information.

Document delivery is an area which, in the past, has seemed to lend itself to basic fill rate and turn-around-time indicators. Relatively few libraries used CAUL Indicator 2. Most libraries seemed to prefer internally developed measures.

Materials availability was, and is, the most used CAUL indicator across Australia. In a large student population, where coursework dominated the environment during the years when physical volumes were critical, this was a very useful measure of student (not exclusively but mostly) frustration. The indicator is no longer for us as central. There is no a priori reason why the indicator couldn't be used in the online environment, and we have tried to get users to respond to e-availability questions as part of the indicator, but we have found that the indicator doesn't reflect a dual reality and thus is becoming more and more marginal each year.

In a survey done as part of the Guidelines work, half of responding university libraries wanted CAUL to publish further indicators to cover electronic resources and their use. Other areas of need identified by that survey included measures of:

-       costing and cost effectiveness
-       targeted client satisfaction
-       information literacy

So how far have we come in the last 5 years?

CAUL has not worked on electronic measures itself, apart from hosting discussions on the topic, but there has been significant progress over the past five years in the global arena, and progress has been surprisingly extensive.

Those electronic measures

Electronic measures have finally arrived, and are now available for use after almost ten years of experiment. The measures are here. They have been described; training materials for collection are starting to become available. Not unexpectedly, the how many (ownership) and how much ($) figures for electronic materials came first. Usage figures have been more difficult and controversial, but even these are now becoming regularised and routine. These e-measures, however, are so far mainly considered optional (read experimental) within official data collection by bodies such as CAUL, ARL and SCONUL. The move for e-measures from optional/experimental/pilot status to mainstream seems to have taken 5 years or more, but the move is almost complete. Working librarians have not necessarily been aware of this progress. Usage figures, until recently, have been the real focus for concern and confusion. CAUL has been using "provision of stock" figures for some years. Pity the staff member who has to complete the electronic serials components of the annual CAUL return. The so-called deemed list is almost routine, but still has challenging aspects. Cost figures in terms of expenditure on e-resources as a proportion of materials expenditure have also made more regular appearances in recent years.

The E-measures we now have available have been standardised and tested over a range of experimental projects for the assessment of electronic library services. In the US, John Bertot has worked with Charles McClure in this field since the mid 1990's. Their list of emetrics consists of 110 measures for electronic libraries produced for the e-Metrics Instructional System and was originally developed for museums and public libraries in the US. These two academics then became part of the ARL E-Metrics Study Team. An e-metrics manual has recently been developed by the pair, together with Andy Lakos of Canada's University of Waterloo. The e-metrics project is slowly proceeding in the ARL way by being tested with a small group of ARL Libraries and then rolled out to more and more libraries.

SCONUL in 2005 introduced a range of new and amended statistics to cover e-metrics, These are based on the University of

Central England 'E-valued Toolkit' of sixteen measures. Charles McClure was the advisor for the E-valued project so these are at least international in scope. The Toolkit measures are grouped under the three areas of:

- Provision of stock
- Use of Services
- Costs

UCE provides training in the collection of the data to UK university staff.

COUNTER

To provide reliable data for the use of e-resources, it has been apparent for some years that the compliance of the publishers is absolutely necessary. Librarians in a number of countries valiantly attempted, for some years, to use their own ingenuity and resources to define consistency in usage across the providers. Libraries across the world were trying hard by 2001 or so to provide e-journal and database usage statistics of a basic sort. Some ARL libraries attempted to analyse, for a defined set of title the basics:

- number of sessions
- number of searches
- number of items viewed/downloaded

Across the Atlantic SCONUL and Joint Information Services Committee also had a project. Even CAUL had their first project running. It was hard work. It soon became apparent that only the publishers themselves could provide the information needed in any sort of standard format. Until that happened, no matter how hard librarians might try comparable figures were elusive.

For librarians in 2005, probably the most spectacular development in this area is the global emergence of the COUNTER[vi] compliant framework over the last three years and its adoption by a huge array of electronic journal publishers.Counter (Counting Online Usage of NeTworked Electronic Resources) began and continues to be based in Britain but has grown from a modest project to an international system endorsed by vendors, industry associations, library consortia and individual libraries. COUNTER recently won an international award in recognition of the exceptional global collaboration between those naturally friendly groups – librarians and journal publishers. It is very impressive how much has been achieved in a couple of short years. The challenge now that COUNTER compliancy has swiftly moved to encompass virtually all significant e-publishers is to use the data as it was intended for e-resource management within libraries.

Like most large libraries UNSW Library had bright individuals who tried to provide reports for four categories of electronic materials:

- databases (for example Medline, of which we had over 95,000 searches in May 2005 alone, or, a less used database World of Science of which we had just on 250,000 searches last year)
- aggregates (such as Proquest)
- publishers collections (such as Science Direct from which there were more than 617,000 full text downloads last year)
- e-mosos (monographs on standing orders for example Routledge encyclopedia of Philosophy or Brittanica Online)

Thanks to COUNTER we have been able to set up monthly and annual reports based essentially on statistics on the numbers of searches and full text downloads. These reports will then appear on our intranet to assist, in particular, those making decisions about value for money prior to renewal time around September. There are some further ratios which can usefully be extrapolated (e.g. $ cost per download). In setting up the reports for technicians to use and compile there is the recognition that the system is all still a little bit clunky in the sense that there is a need for the Library to be proactive. The data only comes on a (virtual) plate when called for. The librarians/technicians must still go to each publisher website when alerted to new statistics, but at least the staff know that, almost all the publishers we have are at least partly COUNTER compliant, which should mean that the data is comparable.

For E-books there is also a draft COUNTER standard so that this happy situation of comparable statistics should soon pertain to all the major e-book publishers. Since e-book collections can often substitute titles based on usage this will be a valuable resource.

The Quantitative in the Qualitative measures

Most qualitative is not separate from the quantitative. Qualitative programs express their top level results in quantitative ways. Qualitative measures can also be reduced to quantitative data for example, the CAUL Rodski indicators and profiles.

Online surveys have given us greater participation rates, ease of administration and much fuller user comments in the free form components of our surveys – it's much easier to type than to write in comments on a paper form. Analysing these can be a guilty manual pleasure or subject to rigorous textual analysis, but it is the quantitative results which when displayed have the most impact. Participation in the Rodski survey across CAUL has been very salutary for UNSW. Having that wide range of data available has provided evidence that many Australian university libraries have a much greater share in the affections of their users than we have. There are many reasons for this. One of the most compelling one for us is demonstrated by the example that our only significant branch library is consistently reaching very high levels of user satisfaction, far higher levels than those shown for our main library. Part of it is excellent staff but part of it is also that very potent 'small is beautiful' factor. Our branch library has poor physical accommodation, but is very atmospheric, complete with balcony, umbrella and resident cat.

Those of you from small libraries should rejoice. Small size is the greatest predictor, all things being more or less equal, I think, of user satisfaction. It is very hard, in my opinion, to love a fourteen storey concrete tower. Rodski overall numbers can be viewed by members of contributing libraries on the CAUL Rodski Indicator Portal. Working through user comments may yield the best benefit, but staring at a graph with the categories of Rodski from your peers etched across the chart above you is truly humbling. Quantitative data wins out, even for the qualitative tools.

Reference & information literacy

From the old reference fill-rate, through WOREP, to CRIG and the ASK model – measures relating to reference have been a minefield. I don't intend to go through this area in detail here, but I do want to mention some commendable work currently being done by the UNISON Reference and Information Services Interest Group on Digital Reference Key Performance Indicators, because this group has a worthy current project which tackles a specific area for e-measures. They (the pioneer group from UNSW, UTS, Macquarie, Sydney and Southern Cross) are proposing a multiple approach based on the ASK model (attributes, staffing, knowledge) to yield both quantitative and qualitative measures for e-reference. Their project is a worthy one and well worth tracking to the conclusion.

The digital reference project approach includes:

- trialling a digital reference client survey
- trialling a survey on staffing and staff training for digital reference
- trialling a digital reference statistical survey

Statistical information being gathered includes:

- service hours open
- service hours staffed
- transactions (reference,non-reference, referrals, follow-ups)
- terminated transactions, time taken, responses turn around time, numbers of users (by category)

Contextual information is also being gathered.

This project is notable because it is cheap, useful, fairly valid, selective, common sense contextual – all my criteria for quantitative data gathering. Watch out for results.

Information Literature/effectiveness of teaching

This is a complex issue to address. Really to get research-based data, your students would have to be split into groups, exposed to a thorough information literacy regime, pinpointed through curriculum mapping and planning and controls and measured on entry and exit to the institution. Some US universities are, theoretically, in a position to do so because of the so-called "capstone" courses which check on key competencies before

graduation. Those US institutions now have common, compulsory programs on entry and exit. There is potential in the US-based SAILS (Standardized Assessment of Information Literacy) project [vii] for a vehicle to assess information literacy skills. SAILS is now in a period of review and even the project staff are not clear exactly what is the appropriate level to use their material (based on meeting the ACRL standards). Is SAILS appropriate for first year students and/or for final year students? Can you really use the same tool at such different levels? There are still more questions than answers. SAILS also relies on quantitative assessment data to report to institutions. There is no avoiding the power of the numbers.

As much as we would like to focus on learning outcomes – what exactly did the students learn? It is sadly true that these sorts of noble professional topics, which have recently been exercising our minds at UNSW, are so easily overcome, even in the university context by the powerful fascination with numbers – with quantitative data. Let me explain.. This year we introduced an online "enabling" tutorial and quiz which has been made mandatory by the Academic Board. We debated at length just what are the entry level skills and knowledge which can reasonably be expected of entering students so that they can undertake their course based information literacy, with confidence. What we had to do in practice to answer that was to develop a tutorial, implement a post-tutorial quiz and introduce a "pass mark" of 80%. What did the academic staff want to know about our progress? Completions rates. Pass rates. Numerical comparisons between Faculties. What we have concluded, after the first 5000 or so students have completed the tutorial and quiz is that what we had produced either more than met, or met, or didn't meet student needs in various measures and that our approach was adored by many and criticised by others. We have more than 2000 sets of evaluative survey data to analyse as I speak. What we did find was that students did this online tutorial from home (85 %), did it very early (mainly prior to session), exactly as we had hoped, and that 5000 students started their program at UNSW knowing a little more about the world of information than they did before. Numbers don't tell it all but they are inescapably fascinating.

[i] CAUL Statistics Website available at www.caul.edu.au
[ii] Rodski Research at www.rodski.com.au
[iii] American Research Libraries at www.arl.org
[iv] Wilson.A., Pitman, L.,Trahn.I, Guidelines for the Application of Best Practice in Australian University Libraries: intranational and international benchmarks. Canberra, DETYA, Evaluations and Investigations Programme, Higher Education Division, 2000. See Publications section of the AARL website for a link to the electronic edition
[v] Wilson, A et al. p.45
[vi] COUNTER at www.counter.org.
[vii] Project SAILS. www.projectsails.org

Biographical note:

Isabella Trahn, B.A. (UQ), Dip. Lib.and M. Lib. (UNSW); worked initially at the National Library of Australia and then for thirty years in university libraries, mainly at UNSW. For the past decade she has had in her portfolio, responsibility for, and an interest in benchmarking and quality issues, statistical reporting and surveys for UNSW Library. She was fortunate enough to undertake an extended study tour in 1998 to research quality management and benchmarking in a series of overseas libraries. In 1999, she was part of the team which produced two publications for DEST on Best Practice in Australian Libraries and then delivered a conference paper at one of the well known Northumbria Conferences on Performance Measurement in the north of England in 1999 on that national project. Currently she is doing many other things besides but still has some involvement with performance measurement. Preparing for this paper was a great opportunity to review some more recent developments in performance measurement.