

SWIN
BUR
NE

SWINBURNE
UNIVERSITY OF
TECHNOLOGY

Why have statistics? ... and what?

Derek Whitehead
CEIRC 2007



Summary of the Session



- 1 The CAUL Statistics Focus Group
- 2 Why stats?
- 3 New Measures
- 4 More issues
- 5 Swinburne thoughts
- 6 Challenges for CAUL

Always start with statistics



A recent survey showed that only 85 significant institutions anywhere had remained virtually unchanged since 1520:

- 70 universities
- Several Swiss cantons
- The parliaments of Iceland, the Isle of Man and the UK
- The Catholic Church
- A few others

Since 979 – Tynwald and the triskelion



What should we count?



CAUL Statistics Focus Group (CSFG)



- Based in Victoria – Cathie Jilovsky, Diane Costello, Stephen Gillespie, Maxine Brodie, Jocelyn Priddey, Gehan Aboud, Christine Wilson (NZ), Janice Van de Velde (NASL)
- Produces annual CAUL statistics
- CAVAL is contracted to collect and publish the data
- Meets about once a year
- There is a network of statistics contacts at each university

CSFG - aspirations



- To publish the statistics as early as possible - consequently, we now tend to publish data as it comes; printed data in AARL will cease
- Provide consistent data over time – standard data and definitions
- Provide an understanding of changes in our environment – but see Paul Genoni in AARL 2004 - “Print serial cancellations in university libraries post 1990: what do the CAUL statistics reveal?”

<http://www.alia.org.au/publishing/aarl/35.3/full.text/genoni.html>

CSFG - ambitions



- Unique contribution to the world of library statistics – the deemed list – we plan to develop a statement on it
- Surveyed users in 2003 and we aim to please:
- Good statistics: clarity, validity, practical usefulness, ease of collection (or already collected)
- Mainly used for: reporting up, analysing, understanding
- Mainly used by: library managers
- Satisfaction: fairly high level of satisfaction

CSFG - adulation



- What do people like about the CAUL statistics? – or any statistics?
 - All in one place
 - The deemed list
 - Quick and easy to use
 - Ability to benchmark (compare) with other institutions
 - Long series of trend data for all institutions
 - Methodology is established and well-understood
 - Rankings and ratios same time

CSFG - ambience



Cheap, useful, fairly valid - why?

- Library statistics are a part of the cost of managing, nothing more – keeping cost/effort down is essential.
- Useful should go without saying, but it is worth saying – if you don't use the data, why keep it?
- Fairly valid is a corrective both ways – you need SOME level of validity, but you can go too far.
- . . . and of course, excessive validity can undermine the usefulness of statistics

CSFG – alterations



- Redevelopment of the statistics site in 2005
 - Based on the 2003 survey & a 2004 proposal by CAVAL
 - Software based on the ARL statistics software
 - New site available Sept 2005
 - New collection methodology
 - See <http://statistics.caul.edu.au>
- Redevelopment of the statistics site in 2005
- Minor redevelopment since then

Why Statistics are Useful



- Determine how well we are going in relation to like institutions and to ourselves, over time
- Understand and improve service delivery
- Provide a basis for resource allocation and budgeting
- What isn't counted isn't valued
- Demonstrate extent and nature of complex changes over time
- Provide a more objective backing for judgement and opinion
- Enhance understanding of customers and their demands
- Opinion/user data provides an objective insight into the customer
- Statistics are an important way to communicate upwards

Why Statistics May Not Be Useful



- They can be extremely time-consuming – libraries put more resources into them than anyone except Finance
- They can be spectacularly inaccurate and meaningless – reference statistics, for example
- They can be extremely misleading
- Comparisons almost always have flaws because of the difficulty of comparing like and like in complex service environments
- Much statistical information collected is NEVER used
- Some opinion/user data can be very unreliable
- Consistency is harder to achieve than you think
- No-one believes them.

What Library Statistics Should Do



Statistics which universities keep and make available could do the following a lot better than they do:

- Give an idea of actual outputs
- Indicate changes in type/balance of outputs
- Relate what we do to what others do – benchmarking
- Understand the customer better
- Help us to plan
- Provide very large numbers to impress
- Change and be relevant but remain consistent

Most faults are trade-offs



- Complex working situation cannot easily be captured by simple statistics – trade-off between usefulness and cost/validity
- Being consistent means closely defining categories and asking respondents to fit the data into these – a cost/validity trade-off
- Collecting only quantitative information is another cost trade-off – but a mantra of management



New measures

- COUNTER development of standard measures
- CONZUL began to collect COUNTER measures
- In 2004 we pilot tested four measures
 - Number of logins
 - Number of database searches
 - Number of full text retrievals
 - Expenditure on online resources

New measures – more



- We added expenditure on electronic resources – but not everyone can provide it
- We tested the COUNTER definition of e-books – early days yet but
 - Are they monographs, subscriptions or a database
 - What about e-books in journal packages
 - What IS a full-text download
 - What is an e-book anyway?
 - Does the COUNTER definition work?

- ththh

New measures – yet more



- We added reciprocal borrowing at a national level – University Libraries Australia
- We discussed e-reserve data – problem is definition, what is in and what not. Do we include lecture powerpoints or not? Do we include links?
- Offshore students – simple enough? Problems include lack of a common definition, what statistics are collected, reliability and standardisation, confusion with distance students.

Other issues – dual sector universities



- Objective is to benchmark dual sector universities with others
- But the data cannot be fully disentangled – only
- Loans data
 - Student numbers – ^{But the} counting is very different
 - Student numbers – but counting is very different
 - Expenditure on collections
 - Customer satisfaction data

Other issues – dual sector universities



- Objective: to benchmark dual sector universities with others
- But only this data can be disentangled
 - Loans data
 - Student numbers – but counted very differently
 - Expenditure on collections
 - Customer satisfaction data
 - Maybe others – info literacy, database use, traffic
- The TAFE deflator is 28 – the Mullarvey deflator is 22

Other issues – customer satisfaction



- Rodski has been used by all universities
- LibQUAL is a rival methodology
- The national benchmarkability impresses people outside libraries
- Rodski has brought
 - Consistent national approach
 - Regular surveys
 - A focus on the customer
 - Cheap, useful and fairly valid

Other issues – use of computers



- Biggest area of complaint in university libraries
- A major focus of service and effort
- ITS can tell us in detail about use
- Comparison of system-generated counts of users with headcounts is very interesting
- So why do regard them as out of scope?

Statistics at Swinburne



- Data driven strategy
- The search for one big statistic
- More measurement
- What are our core KPIs?
- Big numbers impress people
- Is customer opinion the only relevant number?

Data driven strategy



- At Swinburne we have planning and budget drivers which are driven by data
- Two kinds of data should drive strategy
 - Customer data – what they want
 - Output measures – what we produce

Highlighting trends



■ Loans bottoming out?

Year	First-time loans
1999	424,000
2000	401,000
2001	364,000
2002	334,000
2003	316,000
2004	290,324
2005	292,371
2006	317,873

One big statistic



- Loans used to be the big statistic
- And collections. Maybe money.
- Now the statistic is document or item use: book or media loan, full text download, more.
- That is, document use mediated by the library.

“Document use”	Proportion of total
Book & non-book loans	16%
In-house use	8%
Equipment loans	8%
Online reserve downloads	18%
E-book downloads	15%
Journal article full text downloads	36%



We measure more than we ever did

- Loans
- Collections
- Customer stuff
- Info lit
- Traffic
- Space, seats
- Computers
- Opening hours
- Money
- Web statistics
- But not this



Counting Questions



- We used to count all “reference” questions
- We moved to counting all questions
- We classified them into useful categories
- We do this for two weeks and extrapolate
- We have made our methodology widely available

Enquiries – (Total=145,000)	Proportion of total
Directional	11%
Service enquiries	39%
Help with equipment and IT	27%
Information enquiries	19%
Other	4%

Core KPIs



- The key performance indicator is axiomatically good and achievable
- Performance must be outputs
- So which ones are key? Here are mine
 - 1) Item loans
 - 2) Document downloads
 - 3) Traffic
 - 4) Satisfaction



Big numbers impress

Data from 2006

- Catalogue searches 1,093,739
- Online database access sessions 743,827
- Loans (including renewals) 590,387
- Photocopies and prints 1,731,000
- Traffic (turnstile) 1,199,000
- Hawthorn LateLab (after hours) 75,000
- Info enquiries at service desk 159,660

Growth rates impress too



Year	Equipment loans	Growth
2001	5,704	
2002	33,138	481%
2003	62,028	87%
2004	65,924	6%
2005	90,900	38%
2006	116,418	28%

Customer opinion is important



Question: Do you think that students should be allowed to use MOBILE PHONES in the library?

■ Never	1,691	31%
■ Yes, but only in non-silent areas	1,142	21%
■ Yes, but must ring and talk quietly	1,197	22%
■ Yes, any time	1,458	26%
■ TOTAL	5,488	100%

Conclusions, thoughts, challenges



- ❑ A large part of our business is access to computers – what should we count, and how?
- ❑ Equipment loans are the fastest growing use category at Swinburne – should we measure them? How?
- ❑ The library web site is absolutely critical now – what measures do we use?
- ❑ International students are important to us: why don't we count anything about them?
- ❑ Document downloads are now our core outputs measure – but we're not there yet – what has to be done?

Thoughts and challenges



- ❑ Are there other missing statistics? – data we need, don't have?
- ❑ Are the existing consultative and communications mechanisms working?
- ❑ How can CEIRC and the CAUL Statistics Focus Group work together more effectively?
- ❑ What do we really know about long-term trends? Who could do the work?
- ❑ What is going on elsewhere? E.g. Task Force on New Ways of Measuring Collections
<http://www.arl.org/stats/aboutstats/tfnewways.shtml>
- ❑ There is no end to the fascination and usefulness of statistics, even if they confuse at first sight – magic!

Who doesn't love open source?



Who doesn't love Jonathan Creek?



Who doesn't love a challenge?

