
Using survey and usage statistics to improve libraries

Phil Sykes
University Librarian
University of Liverpool
Tel: 0151 794 2673
E-mail: p.sykes@liv.ac.uk

USELESS AND USEFUL STATISTICS

Higher education is awash with statistics; so are university libraries. Some of those statistics are meaningful, providing a decent basis for planning service improvements; others are not. This article is mainly intended as a help to distinguishing between the two. I'm not working to any particularly demanding criterion of usefulness or helpfulness here: in my book, statistics can be useful for library development if they are probably more true than not and help us to exercise judgement in planning or in the allocation of funding.

Perhaps it would be helpful to give an example of a statistic that is useless as a basis for planning service improvements, or at any rate not awfully useful. Suppose you do a survey, in your own university only, which shows that 70% of your users are either satisfied or very satisfied with the library services. At first sight that might appear to be a useful insight, but in reality it probably means virtually nothing other than that your library is not yet bad enough to require UN intervention and that the library staff need not tender their resignation en masse. There are a number of reasons for this. Firstly, although the score *looks* pretty good, because the survey has only been done in your institution there is no real way of knowing whether 70% is laudable or culpable. It could be that if the same survey were to be performed in ten other universities the average score would be 90%, making you a pariah, or 50%, making you a paragon. There are a number of other reasons why your 70% figure probably isn't all that helpful or revealing: it doesn't show what aspects of your library are good or bad; it doesn't show what students think is important; and it may be that the number of students surveyed is so small that the conclusions you try to derive from it are statistically unwarrantable. Section 3,

below, on 'things that make statistics more useful', tries to shed light on some of these issues.

USELESS STATISTICS CAN SOMETIMES BE POLITICALLY USEFUL, THOUGH

Before we plunge into the statistical nitty-gritty I should make a small confession: I have nothing against useless statistics. Sadly, in our fallen world, there are occasions when one has to use them, and I have made abundant use of them myself. A distinction has to be made between using statistics as a basis for library planning and development – what this article is mainly about – and using them politically. Again, let me give an example. Suppose you are Chief Librarian and have acquired a new boss in a management restructuring. Your boss, in the first fine careless rapture of office, has decided that he will interview groups of students to find out what they think about the university. He chooses history students and part-time nursing students as his representative cohorts. They both tell him how awful the library is. He is seriously perturbed. He hauls you in and asks you to account for this sad state of affairs. He gives you fair warning that he may not be able to support your bid for additional resources for the library because his interviews show that it would be throwing good money after bad. What should you do? Warn him that his interview groups were too small to generalise from? Point out that he selected two groups that are always likely to have a more severe view of the library than, say, engineers or scientists? No, because all that, though completely true, would sound too self-justifying and you would not gain your point. What you should do, of course, is to generate just the kind of semi-useless statistics that I described at the beginning of this article. He has played the imaginary king; you have to trump it with the illusory ace. So get going with your general survey, which will show that 70% of students (a thumping majority as you will tell him) are either satisfied or very satisfied with library provision. It will do nothing to help you improve the library, but it will solve your political dilemma.

THINGS THAT MAKE STATISTICS MORE USEFUL FOR SERVICE IMPROVEMENT

Comparability between libraries

My main grouse with the survey I described above – discovering that 70% of students were either satisfied or very satisfied with library provision – was that there was no way of knowing whether 70% was a good or a bad score. If the same survey were to be carried out in a number of

other similar libraries it would make the statistics more meaningful (subject to the tests of statistical significance described below) because you would be able to find out whether your 70% was comparatively good or comparatively bad. This is one of the benefits of the LibQual methodology, used by many UK and hundreds of US libraries, which allows you to compare your LibQual survey results with those of other groups of libraries. Caution does need to be exercised here: the results are probably going to be a bit more useful if the libraries you are comparing yourself with are institutions with a broadly similar mission.¹

One might also argue, of course, that comparing the library with other service departments in the same university would provide useful comparative data. Suppose, for example, your university surveys its academic staff to monitor their satisfaction with different university central services, and the result shows that the library is a more highly regarded service than personnel, estates or finance. Isn't that a useful validation of the quality of your service? Well, theoretically yes, but my suspicion is that the library will tend to be viewed with an intrinsic benevolence and personnel will tend to be judged harshly at *all* universities.² You would really need comparative data with other institutions to derive firm conclusions. For example if the library was massively more popular than personnel at ten similar universities but only slightly more popular than personnel at your institution, then, despite the apparent positivity of the statistic, it would tend to indicate that you had some work to do.

Comparability over time

Statistics obviously become more useful if they can be compared over time, because they give some sense of whether your service is improving or declining. It is important to determine, though, whether changes that take place are statistically significant (see 'Statistical significance', below). It is also important to ensure that circumstances have not changed in a way that means your comparison has become invalid. A good example of this is provided by the SCONUL statistics figures for book issues. These figures include renewals, which, with the advent of OPACs (Online Public Access Catalogues), have become very easy for a user to effect. This means that the issue figures have been hugely bolstered by the explosion of 'touch of a button' online renewals. At Liverpool these now account for nearly 40% of total issues, as against just over 10% in 2000/2001. This means that the apparent buoyancy of issue figures in SCONUL institutions – up from 51 issues per FTE

(Full Time Equivalent) student in 2000/01 to 56 in 2005/6 – is largely illusory. We will go seriously wrong if we use these figures uncritically as the basis of workload planning and service development.

Asking respondents what they think is important

Asking respondents what they think is important is always useful and it's vital if you're using your statistics as a way of deciding on the allocation of resources. Suppose, for example, you have some extra funding that you could put into either extra opening hours or an additional enquiry desk. Your statistics show that satisfaction with opening hours is slightly on the low side, but satisfaction with enquiry support is even lower. If you used this information alone as the basis of your funding decision you would put your money into improving enquiry support. Suppose, however, that you also asked students how important these two factors were to them and discovered that generous opening hours were a lot more important to them than enquiry support. The high importance rating given to opening hours would probably persuade you that you would make the biggest impact on overall student satisfaction by opening the library for longer.

Just as it is useful for libraries to know what their users think is important, so that they can use their resources to best effect, it's useful for universities to know this, for the same reason. So if your university does undertake a satisfaction survey it's worth ensuring, if you can, that importance ratings are included. There is an impeccably altruistic reason for doing this – as a good citizen of your university you would want to ensure that it uses its resources sensibly – but it is also very much in the library's interests. Academic staff and senior managers invariably underestimate the importance of libraries to students. Surveys of what students feel is important are a useful way of countering this, because they invariably show that libraries are more important to students than anything other than the absolutely core academic issues of course content and teaching quality.

Statistical significance

This is an area that was as deeply offputting to me initially, as it will be to the majority of you reading this article. The bad news is that if you really intend to act upon the statistics you gather you do need to know whether they are meaningful enough to form a solid foundation for action. The good news is that, though it sounds nerdily mathematical, statistical significance is largely an embodiment of common sense and there are easy

guides for the mathematically challenged. There are also plenty of people in a university who do understand statistics, so it's reasonably easy to get help if you're out of your depth.

What does testing for statistical significance mean? Let's return to our initial example – the survey in which 70% of the students you surveyed thought that your library was either good or very good. Supposing you repeated the survey a year later and found that, after you had made some improvements to services, the figure was 71%. Could you safely conclude, therefore, that student satisfaction with the library service had improved? To put it another way, could you generalise from the (necessarily limited number of) users who responded to your survey to the whole body of users? You are never going to know for absolutely certain, of course, but what significance testing allows you to do is make a statement like 'It's 95% certain that there has been an improvement' – in other words that the apparent improvement was probably real, not just statistical 'noise'. Without going into the maths of it, the sort of thing that significance levels are influenced by include:

- *the size of the sample:*

Given that what we usually do with student surveys is try and infer things about the student body as a whole from the (typically) minority of students who respond to our survey, it makes sense that we can be more confident about a survey in which large numbers of students participated than one in which only a handful responded. One slightly counter-intuitive fact about samples sizes, though, is that the absolute number of respondents tends to be a more important influence on statistical significance than the proportion of the group that responds. That is why exit polls, asking people how they voted as they come out of the polling station, have proved to be an extremely reliable predictor of general election results, despite the fact that only a very small proportion of the electorate is surveyed.

- *The extent of the difference in scores:*

Again, this accords with common sense. If, in our hypothetical survey, the percentage of students who thought the library was good or very good increased from 70% to 90%, we would expect the rise to be more statistically significant than if it had increased from 70% to 71%. There is, by the way, a useful statistical technique for 'normalising' scores: removing the effect of using different scales for

survey responses (1 to 100 scores as against 1 to 5, for example).

Although all this can all seem aridly theoretical, it is, in reality, urgently practical. For instance, when we undertook our recent LibQual survey at Liverpool, significance testing showed that although we could make meaningful comparisons between the undergraduate responses for 2004 and 2006, the differences in the staff responses between the two years were, except in relation to a small minority of questions, not statistically significant enough for us to make inferences from (which was a pity since they showed an apparent improvement!). Subjecting your statistics to this kind of test also makes them more persuasive, particularly to those academic groups and senior university managers who are statistically savvy and who may initially assume that your statistics are on the fluffy side.

Combining the results of different surveys

No survey you organise, or are involved in, will meet all the criteria listed above. The National Student Survey, for example, offers good comparability between institutions but the fact that it asks only a single, rather badly worded question about libraries means that it offers no insight into what is good or bad about your library. I am a fan of the LibQual methodology, but the level of comparability it offers is limited and it doesn't ask anything about the importance to users of the different dimensions of the library service it investigates. By combining the results from a number of surveys each survey can be made to supplement the deficiencies of the others. Table 1 shows an extract from a table we compiled at Liverpool to create a richer statistical picture of our position by combining the results from the National Student Survey, the OpinionPanel survey, the LibQual survey and an internal Liverpool University survey of student perceptions. This also shows the way in which apparently firm conclusions from one survey can be called into question, or given a radically different perspective, by another that approaches the same issue from a different angle. Our internal University of Liverpool survey, for example, showed a worrying level of dissatisfaction with textbook provision, and might have led us to conclude that recent substantial increases in funding in this area had had no effect. But the LibQual survey showed us that contentment with textbook provision had improved between 2004 and 2006, and that the level of dissatisfaction with textbook provision was now lower than the average for other SCONUL libraries using the LibQual methodology.

Table 1 – Example of combining the results of different user surveys at the University of Liverpool

Statement	Authority
The Library at Liverpool is perceived by students to have improved since 2004.	NSS surveys from 2005, 2006 and 2007. LibQual Liverpool surveys for 2004 and 2006
Student satisfaction with library services is higher than with any other major dimension of student experience at Liverpool (other services, and aspects of course provision such as course content). 81% would recommend the Library to a friend, compared to an average recommendation rate for other aspects of university provision of 56%.	UoL Student Experience Survey 2007
Student satisfaction with our library, as reflected in the National Student Survey, has improved relative to the satisfaction of students at other civic universities with their libraries, though we are still behind four of the eight comparators.	NSS surveys from 2005, 2006 and 2007
BUT when the statement on which students are asked to comment emphasises 'opening hours', as the OpinionPanel survey does, we have among the top four satisfaction rates in the country	OpinionPanel survey 2007
The area of library operation that has improved most since 2004 is staff support (though access to subject specialist help is still fractionally below the minimum acceptable level)	LibQual Liverpool survey 2004 and 2006
The least improved area since 2004 is library buildings and environment, though the creation of the informal study area on the ground floor of the Harold Cohen Library appears to have improved the results a little between 2004 and 2006	LibQual Liverpool survey 2004 and 2006
AND the only areas where we are perceived to be below the average for libraries using the LibQual methodology relates to buildings and study environment	LibQual UK norms for 2006
The level of dissatisfaction with library stock has decreased substantially since 2004	LibQual Liverpool survey 2004 and 2006
AND satisfaction with library stock at Liverpool is higher than the average for the other UK university libraries using the LibQual methodology	LibQual UK norms for 2006
<p>BUT</p> <ul style="list-style-type: none"> • provision of 'the main texts and and readings I need for my work' is still, even at the improved 2006 level, considered by students to be below the minimum acceptable level (though when the statement to which students respond is phrased as 'the printed materials I need for my work' students place provision just above the minimum acceptable level) • 27% of students are either 'not satisfied' or 'not at all satisfied' with the Library's provision of the books they need for their course • The free text responses to the National Student Survey reflect strong discontent with textbook provision 	<p>LibQual Liverpool survey 2006</p> <p>UoL Student Experience Survey 2007</p> <p>National Student Survey 2007 – free text comments</p>

Between opening hours, staff support and stock provision the highest level of satisfaction is with opening hours and the lowest is with stock provision	UoL Student Experience Survey 2007
AND levels of satisfaction with book provision are much lower than with journals or electronic information provision	UoL Student Experience Survey 2007 LibQual Liverpool survey 2006
Students regard the Library as more important than any other dimension of their university experience apart from core course attributes such as course content and organisation. 91% of students regarded the Library as important or very important. The equivalent average score for other aspects of university provision was 56%	UoL Student Experience Survey 2007

Table 1 also provides an illustration of another issue that should be borne in mind when using survey results to make planning decisions: that small changes in the wording of a question can make an enormous difference to users' responses. There were two surveys in the period covered in the chart that essentially aimed to elicit how good the students thought the University of Liverpool library was. In one, the National Student Survey, we were somewhere about the middle of the field. In the other, the Opinion Panel survey, we were third in the Russell Group, with only Oxford and Cambridge above us. The clue to the apparent paradox lies in the wording of the questions asked. In the NSS students were asked to respond to the statement that 'The library resources and services are good enough for any needs', while the statement to which students were asked to give their response in the Opinion Panel survey was 'Good library and library opening hours'. The clue to our superior performance in the latter survey was probably the fact that we had recently introduced 24-hour opening – a fact that would have been brought to the forefront of people's minds by the specific reference to opening hours in the Opinion Panel survey.

Asking respondents what they think is an acceptable level of service

Most of the satisfaction surveys that libraries conduct simply ask respondents how good they think our services are. LibQual asks them, in addition to this, what the minimum acceptable level of service is and what the desired level of service is. This is a powerful technique that can fundamentally alter one's interpretation of satisfaction data. Let me give an example. In our LibQual surveys for 2004 and 2006, at Liverpool the biggest single improvement came when undergraduates were asked how satisfied they were with 'the main texts and readings for my work'. This might have prompted us to turn our attention to other areas

where the level of satisfaction had increased less or had stalled. BUT, because we had asked students to define the minimum acceptable level for this factor we knew that even the dramatically improved level of provision was still below the minimum acceptable level. Similarly, the LibQual results showed that postgraduates perceived the library service in 2006 to be better than in 2004. *However*, it also showed that their perception of what was an acceptable minimum level of service had become more demanding, so that the extent to which they perceived the library to be better than the minimum acceptable standard had, in fact, declined. So we knew we still had work to do on postgraduate satisfaction.

Using statistics at the right level of detail

Probably the biggest obstacle we face to making use of much of the survey data we collect to help us improve services is simply that it is too general. Even if you know that, for example, your library's score in the National Student Survey is poor relative to other similar universities, you cannot do much about the problems it indicates without more detailed data.

I have no solution to this problem, but there are perhaps some things we could do collectively to help ourselves. One might be to pool our efforts in the analysis of big datasets to extract the data we need, or for SCONUL and/or RLUK (Research Libraries UK) to sub-contract the work on our behalf. One obvious target for this might be the National Student Survey statistics. I said above that the general score for your library in the NSS is of little help to you in improving services. There is, however, a wealth of data at subject level that might usefully inform your strategic planning by revealing areas of particular weakness. The reason few of us make use of it is that it is so difficult and laborious to extract the relevant data from the datasets available to us.

Straying a little from survey data we could also, collectively, bring pressure on LMS (Library Management System) and other software suppliers to provide better – or more readily accessible and intelligible – management information. Consider how amazing it is that, in a typical library survey, we ask people how often they visit the library. It's amazing because that data is already there, 100% accurate, in our systems. It's a bit like a bank basing its accounting on asking its customers how much money they think they have in their savings accounts. We don't extract the data because it's simply too fiddly and cumbersome to do so.

Even very general surveys will, if they contain a section for free text comments, of course yield data that is sufficiently detailed to be of use for planning purposes. Again, the National Student Survey is a case in point. The free text comments for your university are made available to it, so ask to see them if you haven't received them.

IS IT WORTH PUTTING THE EFFORT INTO ANALYSING SURVEY STATISTICS?

I can sympathise with a degree of cynicism about survey statistics. I think this is partly engendered by the way we and our universities manipulate them for competitive advantage. Competition between our universities has, in some ways been a great galvanising force but, as we all know, the first casualty of markets is truth. Our universities are quite as shameless about flaunting statistics that show them in a good light as they are ruthless at suppressing those that show them in a bad one.

Even those who are after truth, rather than competitive advantage, can be discouraged. It can take a lot of work, and a number of overlapping surveys before you get anywhere near 'the truth'. And often that truth is quite banal: because we are a reasonably empathic band we already have a pretty good idea of what our users like and don't like, and of the ways in which our services fall short of the ideal.

However, there really are good reasons to persevere and these are likely to grow more salient as the years pass. One reason is that competition between universities to recruit, and to recruit the best students, will grow and become more global. In such circumstances it becomes commensurately more important that we have an accurate picture of what distresses and what delights our students. That is why private-sector firms spend so much money on market research. But, more

fundamentally, knowledge of our users' needs and preferences should be seen as a core professional competence. We are all used to people talking nonsense about libraries, but we need to be able to refute the nonsense with convincing and rigorous data. That knowledge is also a powerful weapon in our dealings with our own universities. I have tried to avoid writing about the 'political' uses of survey statistics. But there really is nothing as good for your budget prospects as, for example, being able to show that investment in the library has led to improved student satisfaction.

NOTES

- 1 It's also worth noting – although it is outside the scope of this article – that if you are using the comparative statistics politically (say, to buttress a case for additional funding) they will be more persuasive if the comparison is with universities that are perceived by your university's senior managers to be similar. There is little political capital for me, at Liverpool University, for example, in putting forward a case based on the idea that our funding is inferior to that of Oxbridge or the Ivy League.
- 2 Having said that, I would find it difficult to resist making meretricious political capital out of such a finding!