Assessing library web page usability: how benchmarking can help

Sue White

Author

Sue White is Deputy Director of Computing and Library Services at the University of Huddersfield and led the benchmarking project described in this article.
E-mail: s.a.white@hud.ac.uk

Abstract

This article looks at how the 'Mystery Shopper' methodology used by a consortium of universities was adapted to assess the usability of the library web pages of all four universities. Using a methodology refined from a previous web usability exercise, the various elements in the process are described. The outcomes of the project are briefly discussed, along with lessons learnt from the process.

Acknowledgements

Introduction

A benchmarking consortium comprising four universities has been working together since early 2000, and each year since then has undertaken three or four projects on mutually agreed areas (Hart, 2002). The universities involved are Derby, Huddersfield, Leeds Metropolitan, and Staffordshire.

The subject of this article is a project undertaken in 2003, to assess ease of access to information on, or via, the Library / Learning Centre web sites. The aim was to identify any problems with usability, to share good practice, and to make recommendations for improvements. The article focuses primarily on the methodology used and the lessons learnt.

'Mystery shopper' methodology

We refined the methodology used in our project of the previous year, which included a 'mystery shopper' exercise to assess usability of web pages by off-campus users (Creaser, 2003).

Our version of 'mystery shopper' was an adaptation of the conventional usage of this methodology. Rather than posing as customers and testing individual staff members' responses to questions, our 'shoppers' were asked to test each partner institution's web site against a pre-determined exercise sheet – essentially to undertake usability testing.

Our approach was in part informed by the literature on usability testing (e.g. Nielsen, 2000) but essentially we took a pragmatic view of what could be achieved within the timescale, and planned the project accordingly.

'Shoppers'

Each partner institution secured the services of four 'shoppers', or volunteers, in the following categories
- Undergraduate student
- Postgraduate student
- Academic staff
- Dyslexic student.

We hoped these categories would give us a reasonable cross-section of library users, ranging from the relatively inexperienced to the – in theory – relatively advanced. We particularly wanted to include dyslexic students in the study, given the large numbers of dyslexic students at each of our institutions and the need to comply with new DDA legislation.

Securing the services of these sixteen shoppers was no mean task, particularly with regards to academic staff. Considerable persuasion was needed in some cases, and the inducement of £10 book tokens and refreshments certainly helped. We approached dyslexic students either directly, through library contacts, or via Student Services.

Role of the Observers

Crucial to the success of the exercise was the role of the 'Observers'. We decided that two
Observers should be present at each trial, one to record the routes taken when navigating through the site (complemented by web logs where available), and the other to record the time taken to answer each question, and any technical problems which would impact on the shopper's ability to find the required information. Prior to the commencement of each trial, the Observers explained the purpose of the exercise, reassuring each shopper that we were testing the usability of the web site and not their competence as a web-searcher. During the trial the Observers intervened as little as possible, except to supply usernames and passwords at the appropriate time. The Observers also administered the Pre-questionnaire Diagnostic, and the Exit Interview.

The project group generally favoured using two Observers, because it shared out the various duties required and relieved the responsibility which could be felt by a single person. One or two group members however, thought some shoppers might have felt intimidated by having two people in the room with them.

Pre-questionnaire Diagnostic

In order to gain an indication of the web skills of each of our shoppers, we devised a simple pre-questionnaire diagnostic. This comprised a brief résumé of each shopper (institution, subjects studied, form of dyslexia) as well as four questions to ascertain their experience of Internet use (Figure 1):

1. How often do you use the internet?	Daily	Weekly	Monthly	Less often	Never
2. What do you use the internet for?	Leisure/ entertainment	Study	Research		
3. How often do you use the University web site?	Daily	Weekly	Monthly	Less often	Never
4. How often do you use the Learning Centre web pages?	Daily	Weekly	Monthly	Less often	Never

Figure 1: Diagnostic

Although a somewhat crude device, it did give us basic information against which to analyse the results. It also dispelled some of our preconceptions that – for example – postgraduates and academic staff would be more sophisticated Internet users than first year undergraduates.

The questionnaire

Building on our experiences of the previous year's benchmarking project, we reduced the number of questions to six, and ensured each one of them tested a specific aspect of the web pages.

The questions were:

1. From the University home page, find the Library/Learning Centre pages
2. Check the library catalogue to see if the Library/Learning Centre has copies of Black skin, white masks by Frantz Fanon.
3. Find the weekend opening hours of the Library/Learning Centre during term time
4. Find an example of a web site which we recommend to be a good example for information on Law.
5. a) How would you obtain the password for Cinahl (a Nursing database)?
   b) Log in to Cinahl
   c) If you were having any problems accessing Cinahl (or any other electronic resource), where would you go for help?
6. Which years of the British Journal of Psychology are available electronically in full text?

Each question was assigned a time limit, beyond which it was deemed the site had failed to be sufficiently usable. We estimated the questionnaire would take approximately half an hour to complete.

The questions were standardised, so that each shopper worked through an identical questionnaire for each library web site. We felt this was the best way to secure comparative data, but in retrospect it would have been more sensible to customise according to institution. Question 3 for example caused confusion, because multi-site libraries have site-specific opening hours and we did not specify which site was being referred to.

Most shoppers decided to do all four web trials in one session, though in some cases with a break for coffee half way through. For the less experienced web users who took longer to complete the trial, it was a tiring experience. In any future projects we would have to consider either simplifying the questions (which then risks not being a sufficiently rigorous test of the web pages), or allowing more time for the trial (which may make it even harder to recruit volunteers).

Navigation routes

Prior to the trials, partner institutions supplied 'preferred web routes' for each question (an example for question 4 would be: At Quick Links click on Subject Information; at Subject Resources by School click on Business and Law; click on Law; click on Websites). Observers therefore had a set of preferred routes in front of them when conducting the trials. During the trial they recorded the actual routes taken, and were also able to observe the search behaviour of each shopper. This behaviour included things like whether or not the shopper used the navigation column or scrolled down the page, which part of the screen they naturally focused their attention on, and what they did when they got lost.

Some institutions supplemented the Observers' records of routes taken, with computer-generated web logs. These provided an accurate picture of the pages accessed, but were less useful in determining search behaviour than the manual process employed by Observers.

The exit interview

Like the Pre-questionnaire Diagnostic, this was brief. We were conscious that the shoppers had given up their time to participate in the project, and did not wish to draw out proceedings any longer than necessary. The questions were:

1. What do you think of the look of the site?
2. How easy was it to use?
3. What did you like about the site?
4. What did you dislike or find annoying about the site?
5. Any other comments.

For each question we asked for comments, and additionally for questions 1 and 2 asked them to rate the web site on a scale of 0 to 5.

The results: question by question

Whilst it is not the purpose of this article to report the detailed results of the project, readers may be interested in a brief summary. The failure rate recorded in the end columns below were due either to the shopper not completing the question within the allotted time, or due to technical problems.

Question 1. From the University home page, find the Library/Learning Centre pages

Average response time	Shortest response time	Longest response time	Failure rate
41 secs	1 sec	1 mins 57 secs	0%

As expected, those sites with a direct link to Library pages from the University's home page were accessed significantly faster than those without.

Question 2. Check the library catalogue to see if the Library/Learning Centre has copies of 'Black Skin, White Masks' by Frantz Fanon.

Average response time	Shortest response time	Longest response time	Failure rate
1 min 51 secs	4 secs	4 mins 32 secs	14%

One of the biggest reported problems was in getting back from the catalogue to the library home page having completed the search. Some shoppers knew they had to leave the catalogue and return to the web pages but could not, whilst others appeared not to realise they had to leave the catalogue in order to answer the next question.

Question 3: Find the weekend opening hours of the library during term time

Average response time	Shortest response time	Longest response time	Failure rate
1min 31 secs	14 secs	4 mins 55 secs	11%

Most shoppers answered this question within the time allowed (5 minutes), but with wide variations. The sites which performed best were those with a clear link from the library home page.

Question 4. Find an example of a web site which we recommend to be a good example for information on Law.

Average response time	Shortest response time	Longest response time	Failure rate
2 mins 24 secs	22 secs	4 mins 59secs	30%

This question was designed to test how easily users could locate the subject pages, and whether they could use them effectively to locate quality approved information sources. Because of the ambiguity in the phrasing of the question, shoppers adopted a variety of approaches to answering it and many did not use the subject pages. Several observers noted that the shoppers shunned the more traditional hierarchical structure of library web pages, preferring instead to use a single search box. Some shoppers did not seem to know what they were looking for, and even when they did find the subject pages on law, were unable to recognise an appropriate link.

Question 5: a) How would you obtain the password to Cinahl b) log in to Cinahl c) If you were having problems accessing Cinahl (or any other electronic resource) where would you go for help?

	Average response time	Shortest response time	Longest response time	Failure rate
5a	1 min 12 secs	20 secs	9 mins 35 secs	19%
5b	1 min 14 secs	4 secs	5 mins 32 secs	28%

5c	1 minute 3 secs	5 secs	4 minutes 50 secs	28%

This question aimed to test ease of access to a subscription–based electronic information resource, requiring a password. 81% of shoppers successfully located the source of the password, but logging on still proved problematic.

Question 6: Which years of the *British Journal of Psychology* are available electronically in full text?

Average response time	*Shortest response time*	*Longest response time*	*Failure rate*
2 minutes 25 secs	37 secs	6 minutes 41secs	34%

This question had the lowest success rate of all questions, with 34% failures, half of which were due to technical problems. Many shoppers tried to get the information from the catalogue first, but not all catalogues contained e-journals. This raised the interesting question of where users instinctively look for information (which must in part be based on the experience of using their home institution's web pages), but it should also be noted that the degree of prominence assigned to e-journals on the web pages determined the ease with which users found them.

Feedback from shoppers

The shoppers rated each site according to 'Look' and 'Ease of Use' on a scale of 0 to 5 where 5 was best. The average score for 'Look' was 3.05, and for 'Ease of Use' 2.93, and whilst this masks variations between sites, it shows that none of the sites was outstanding.

The qualitative feedback was interesting, but often contradictory. One site attracted the following comments:

'Like the fact there's lots of options, the alternatives make it easy'
'Sometimes confusing – too many options'

Two contrasting comments on another site were:

'Too much clutter'
 'Very easy. Found the information quite quickly.'

As expected, the shoppers tended to rate their home sites more favourably, presumably demonstrating the familiarity factor.

What we learnt from the project: improving our web pages

We derived three valuable sets of data – the quantitative (time taken, record of failures, rankings assigned to sites at the exit interview), qualitative (comments from shoppers on each site), and the record of the routes taken when navigating the web sites. The latter gave a rudimentary insight into user searching behaviour, and has helped inform at least some of us about appropriate navigation and web architecture for the future.

Each partner institution has acted on the data as it sees fit, but some general principles of good practice have emerged from the project. These are similar to findings of other usability exercises:

- Keep navigation simple and ensure there is an obvious route back to the library home page.
- Provide multiple access routes, to cater for different searching behaviour
- Avoid too much text, or 'clutter' on web pages. This is especially important for dyslexic users.
- Cut down on jargon. Do not use terms which may not be understood by the layman and avoid terms which could be ambiguous (e.g. 'electronic resources').
- Do not make assumptions about the skill levels of users. Some dyslexic users, for example, were considerably more proficient than academic staff.

The following are particularly pertinent in the light of the findings of projects such as JUSTEIS and JUBILEE:

- Basic information literacy skills are essential to enable the most effective use of web pages and the information resources to which they link.
- Hierarchical subject pages appear to be redundant for many library users. The internet-net type search box is the favoured vehicle for information searching.

What we learnt from the project: the methodology

The methodology generally worked well. We had some initial problems in securing the volunteer shoppers, and in retrospect we were overly optimistic in thinking that each shopper's input could be completed within two hours. Next time we would pilot the questionnaire to get a more realistic idea of the time needed.

In terms of staff time, the project did require considerable commitment from two or three people in each institution. For the majority of people on the group, the project was undertaken in and amongst their day jobs. The work was essentially divided into three phases:

Phase 1. Preparatory work by the whole group to establish the process, to agree what data to collect and how to analyse it.

Phase 2. The mystery shopper trials, which took place over a two-week period. This was a particularly intensive time for the Observers.

Phase 3. Analysis of the data and writing up the findings, which took several more weeks.

On the positive side, we all enjoyed the experience of working together, and felt we achieved something really worthwhile at the end of it. As a result we have decided to continue our work in the current year, focusing in more depth on dyslexic and visually impaired users.

References

Creaser, C ed. (2003). As others see us: benchmarking in practice, Loughborough, Library and Information Statistics Unit (LISU) Occasional Paper 33, ISBN 1901786536
Also available at: http://www.lboro.ac.uk/department/dis/lisu/lisuhp.html [24.02.04]

Hart, L. (1992). Benchmarking consortia: working together in practice, *SCONUL Newsletter* 22 Spring 2001, 17-18

JUBILEE JISC User Behaviour in Information Seeking: Longitudinal Evaluation of EIS
URL:
http://online.unn.ac.uk/faculties/art/information_studies/imri/rarea/im/hfe/jub/hfjubile e.htm
[25.02.04]

JUSTEIS JISC User Surveys: Trends in Electronic Information Services.
URL: http://www.dil.aber.ac.uk/dils/research/justeis/cyc2rep.pdf [25.02.04]

Nielsen, J. (2000). Why you only need to test with 5 users. URL:
http://www.useit.com/alertbox/20000319.html [26.2.04]
See also numerous other articles by this author at http://www.useit.com/alertbox

WORD COUNT: 2846