NHS Wales e-library portal evaluation (For Informing Healthcare Strategy implementation programme)

Thomas, Rhian; Urquhart, Christine

Publication date:
2006

Citation for published version (APA):
Thomas, R., & Urquhart, C. (2006). *NHS Wales e-library portal evaluation (For Informing Healthcare Strategy implementation programme)*. National Health Service. http://hdl.handle.net/2160/217

NHS Wales e-library portal evaluation

Report for

Informing Healthcare Strategy
Implementation Programme

By

Department of Information Studies
University of Wales Aberystwyth

Authors: Rhian Thomas, Christine Urquhart

May 2006

Contact details
Dr Christine Urquhart
Department of Information Studies
University of Wales Aberystwyth
SY23 3AS
Tel. 01970-622162, fax 01970-622190, email cju@aber.ac.uk

Acknowledgements

The research team are grateful to all the health librarians who contributed to the project. The knowledge gained from the workshops was useful in planning the format of the evaluation work and we are particularly grateful to the librarians who made arrangements at their sites for volunteers to attend the evaluations. We are very grateful to the health staff who gave up their time to take part in the think aloud sessions.

List of abbreviations

CINAHL: Cumulated Index to Nursing and Allied Health

CPD: Continuing Professional Development

HOWIS: Health of Wales Information Service

IT Information Technology

NICE: National Institute for Health and Clinical Excellence

Table of contents

Executive summary

The main messages arising from the think aloud sessions are listed below. Sites involved were Ysbyty Gwynedd, Bangor; Ysbyty Glan Clwyd, Bodelwyddan; Wrexham Medical Institute, Singleton Hospital, Swansea and Llandough, Cardiff.

Staff taking part in the sessions included Clinical Researchers, Specialist Nurses, Physiotherapists, R&D Officer, Clinical Audit Officer, Doctors, Dietician, Biomedical Scientist, Pharmacist, Medical students, Clinical Biochemist, Senior Manager, Radiographer, Mental Health nurse.

MAIN POINTS:

1. Search engine box at the top should be removed – or improved. The main problem arising is that users DO NOT UNDERSTAND that it only searches on that particular page and does not carry out a general search across the site for their request. There are instances where people only get that far and give up thinking there's nothing on the site for them.

2. MEDLINE needs to be named. Currently, only found under OVID DATABASES. Most people we saw did not know Medline was part of OVID and did not look under OVID in order to find Medline – it needs to be more prominent.

3. Tabs at the top of the page need to be larger. Most people ignored them and immediately went to use the left-hand side short-cuts. Again, tabs need to be more prominent. After they were pointed out, they said they were useful – but didn't see them at first.

4. What about a FAQs tab? Need a quick navigation through the site to point users in the right direction – particularly novice users who seemingly cannot take in everything on the page.

5. Lack of searching skills issues were evident – one doctor had great trouble getting to grips with an online textbook - consistently did not realise they could scroll down the index/chapters to find what they wanted – they would have given up using it had they not been shown at the time.

6. Lack of awareness issues also evident. Many of those seen were unaware of HOWIS and the e-library. There is an issue with Trust layout on the separate Intranet sites and a lot of people accessed resources directly from the Trust links on the front page.

1 Introduction

1.1 Aims and objectives

The findings of the Wales User Needs and Knowledgebase study have implications for the design of the e-library portal, in particular the needs of users from staff groups that have not traditionally been NHS health library users. The findings suggest that workshops are necessary to help in user-centred design activities.

The objectives of the extension work were to:

- Prepare briefing materials for the workshops (n=2) for the library managers, assist them in carrying out their own usability investigations (to extend the workshop activities, particularly for the staff groups who may be unable to attend the workshops)
- Organise four workshops (North, Mid, SW and South Wales) for e-library users, facilitated by some library managers, to demonstrate the usability assessment methods, and to gain as much information about the major usability criteria and design considerations for users
- Collate information gained from work done in primary care

Briefing materials were to be provided at the workshops. The Wales User Needs and Knowledgebase Tools study indicated that library staff had little contact with some of the staff groups who need to be served by the e-library. The aim for the usability assessment approach proposed was to help initiate discussions with some staff groups as well as increasing informed library involvement with the HOWIS e-library project.

1.2 Scope

The following usability techniques were reviewed for the briefing materials, with instructions provided on 'how to do this' as well as details of any published and relevant case studies.

- Think aloud protocols (using scenarios for users to work through)
- Heuristic evaluation (using recognised criteria)
- Prototype/scenario research using mock-ups to assess (mainly) navigation and functionality
- Card sorting exercises (to assess vocabulary problems)

1.3 Revisions to workshop plans

As the workshops could not be co-ordinated easily with training sessions, more emphasis was placed on 'think aloud' sessions conducted with individual users. This made organisation of the usability sessions easier, and the librarians had recommended focusing on this type of evaluation in their workshops.

2 Methods

The approach taken was to:

- Review literature on usability testing, to assess the feasibility and applicability of various techniques
- Discuss the briefing paper on usability testing at workshops with librarians (two workshops were held on 9 February 2006– South Wales, and 21 February 2006 – North Wales)
- Conduct usability assessments with individual users at the sites agreed (three in North Wales, two in South Wales)
- Observe one training session

3 Results

3.1 Briefing paper

The briefing paper (Appendix one) reviewed the aims, advantages and disadvantages of:

- User protocols – normally think aloud protocols
- Heuristic evaluation
- Paper prototypes and scenarios
- Card sorting tests

Literature on the use of each of these techniques was identified, with emphasis placed on applications in the health sector.

3.2 Workshops

The librarians attending the workshops discussed the briefing paper and concluded that:

- Card sorting tests, though interesting, involved a fairly sophisticated knowledge of the content of the named resources
- Heuristic evaluation probably better for later stages
- Paper prototypes and scenarios – possible (but hard to track)
- Think aloud sessions were likely to be more productive for novice users of NHS Wales e-library portal.

3.3 Usability assessments

Usability assessments were conducted at three sites in North Wales (Bangor, Glan Clwyd, Wrexham) and two sites in South Wales (Swansea and Cardiff).

The staff involved included a wide range of staff:
Clinical Researcher, Specialist Nurse, Physiotherapist, R&D Officer, Clinical Audit Officer, Doctor, Dietician, Biomedical Scientist, Pharmacist, Medical students, Clinical Biochemist, Senior Manager, Radiographer, Mental Health nurse.

The usage of the NHS Wales e-library and e-resources varied considerably.

In total, 19 think aloud sessions were obtained from users, together with observations made at one training sessions.

3.4 Usability issues

3.4.1 Search engine box

Users appear to expect that the search engine box at the top will search the entire site and everything in the various parts of the site. Searches are frequently for specific clinical aspects e.g. point of care testing, complications of diabetes. When nothing is found, with no explanation, the immediate reaction is puzzlement and many simply repeat the search, maybe modifying the search terms slightly but rarely do they use terms that are general enough to indicate that there are types of resources to help them. For example, while the searcher wants a short summary of the complications of diabetes, terms such as reference book are not included in the search string. There are instances where people only get that far and give up thinking that there is nothing on the site for them. . The main problem arising is that users do not understand that it only searches on the ‘front pages’ and there are no error messages that indicate why they are getting zero results.

Recommendations
Search engine box at the top should be removed – or improved, with at least a link to FAQs or explanations of common reasons for a null result.

3.4.2 Finding MEDLINE

Most of the searchers knew about MEDLINE, but only the more experienced searchers knew that MEDLINE was part of the OVID databases. A high proportion did not find MEDLINE (until it was pointed out to them at the end of the session).

Recommendation
MEDLINE needs to be named as a link, to make it more prominent, and to lead people into other possible databases they might use.

3.4.3 Tabs for navigation

Most people ignored the top tabs and immediately went to use the left-hand side short-cuts. And that has the problem in that searchers tend not to scroll down. If they rely on the left hand side and don't scroll down, they are likely to miss links to resources. After the top tabs were pointed out, searchers said they were useful, and could think of occasions when they might want to use the resources under the various headings.

Recommendation
Tabs at the top of the page need to be in larger font.

3.4.4 Frequently asked questions

There were several common unsuccessful strategies used. Not all of these can be attributed to the features of the current site design, as some are problems inherent to the navigation features of the resources.

Common problems encountered were:

- Locating guidelines successfully
    - *searchers missed the subject required as the list was lengthy*
    - *some confusion over the application of the guidelines to the Trust – were these local, NHS Wales or just England or anywhere else?*
    - *searchers missed search boxes within the guideline resources*
    - *guideline compilations not always labelled as users expect – technology appraisals imply equipment to many people*
    - *several searchers commented that they would expect to find NICE guidelines by Googling 'NICE' and then searching directly on NICE pages. They were therefore not quite sure what they were getting on the NHS Wales e-library*
- Not realising that resources were two clicks away
    - ***Advantage** of the description above the resource link was that searchers could assess whether the resource was likely to be useful*
    - ***Disadvantage** was that some searchers didn't see the resource click, and instead resorted to the search box (unsuccessfully).*
- Searching for journals was hit and miss

- *Searchers did not understand how to search within a journal collection or within a particular title (OVID Main search page not found)*
- *Their interpretations of 'library (local) holdings' varied but most were unclear about the meaning and implications for access*
- *Users accustomed to PubMed, and university access, tend to access journals directly from the links in PubMed and do not realise that it is only their university setting this up for them that makes the access seamless*
- *E-journal listings seem to look different depending on the route used to get there, and users are not sure whether they are searching the NHS Wales subscriptions, all possible OVID e-journals or the subscriptions available to them locally (NHS Wales plus local collections)*
- *Some searchers have journal sites already set up in their 'favourites', and a few who are more involved in research are set up to receive table of contents information. They would expect to be able to click through to the full text or download the pdf as a fairly painless operation*
- *Athens authentication can 'get in the way' when doing searches for journal articles*

- Not scrolling down, to find more information. Searchers, particularly novice ones, simply did not seem to realise that there might be more at the bottom of the page, including a link or two.

Recommendation
A FAQs tab would be useful to give some quick advice to those not finding anything or finding too much. Librarians should be able to provide further details to help compile a list of the top ten searching problems, or to verify that the above list fits in with their perceptions from training sessions.

Resource descriptions are useful to the novice searchers to confirm that they are about to search the right type of resource but the link needs to be made more prominent, at the top, perhaps, above the description so that the more experienced users can click through directly, without the need to scan the text to find the link.

3.4.5 Searching skills

Boolean searching strategies are alien to those used to Google. Beyond a simple 'and' with two terms, mistakes are common. A common error is to assume that X AND Y OR Z means X and (Y OR Z).

Navigational and searching skills were often lacking – one doctor had great trouble getting to grips with an online textbook - consistently did not realise they could scroll down the index/chapters to find what they wanted – they would have given up using it had they not been shown at the time.

Medical students commented on the need to be aware of the type of sites that patients may have used, and that did mean that health professionals need to keep an eye on 'Hitting the Headlines' and some of the patient support group sites in their speciality.

Recommendation
Health librarians in Wales could pool resources to produce 'Ten things you should know about searching the NHS Wales e-library', with very basic instruction on Boolean searching with 'AND', 'OR' operations, the basics of using the e-books, things to remember about finding guidelines, and use of Bandolier (for example), or other resources that provide quality information for patients and professionals.

3.4.6 Lack of awareness

Lack of awareness was evident. Many of those seen were unaware of HOWIS and the NHS Wales e-library (or that there was a difference). There is an issue with Trust layout on the separate Intranet sites and a lot of people accessed resources directly from the Trust links on the front page. It cannot be stressed enough that unless the resource behaves like a search engine most searchers are unsure how to go about the search.

Recommendation

There has been promotion – but this is obviously an ongoing task. As nurses were aware of the RCN site and its e-resources, perhaps working through professional groups is, or will continue to be the most productive way of ensuring that staff within Wales are aware of the resources that are available for them.

Appendix 2

USABILITY TESTING OF E-LIBRARY PORTALS

BRIEFING PAPER

Prepared for A2K team Informing Healthcare Care Programme, NHS Health Libraries in Wales

By

Department of Information Studies
University of Wales Aberystwyth

Contact author

Dr Christine Urquhart
Department of Information Studies
University of Wales Aberystwyth
Llanbadarn Campus
Aberystywyth SY23 3AS

Email cju@aber.ac.uk, tel. 01970-622162

1 Introduction

1.1 Aims and objectives

The aim of this briefing paper is to provide guidance on usability techniques that health librarians could apply in their own library settings, to inform the future development of the A2K e-library portal. There are several usability techniques that require special software or equipment but details of these are excluded from the scope of this study.
The objectives are to:

- Identify appropriate usability testing techniques
- Describe how these are applied
- Discuss any case study applications, noting the advantages, and disadvantages of particular techniques
- Provide details of further reading for follow-up.

1.2 Usability testing for library Web sites

Usability studies are often conducted alongside other user needs studies, as is the case for the A2K e-library portal. A survey conducted by the Digital Library Federation (USA)[1] found that member libraries frequently conducted user surveys (questionnaires), and focus groups to gather general satisfaction data, but notes that interface design questions are hard to answer from questionnaire data.

The common methods used for usability testing are:

- User protocols – normally think aloud protocols
- Heuristic evaluation
- Paper prototypes and scenarios
- Card sorting tests
- Transaction logging and remote monitoring (not discussed in detail)

Each of these is discussed in the following sections. The same framework is used for each technique:

- What the method aims to achieve
- How it works
- Advantages
- Disadvantages
- Examples of its usage
- Further references

There is some debate about how easy e-library portal use should be or could be. For example, Nicholas Joint[2], defines the portal as a searchable network information retrieval service powered by an effective search engine that gives access to content, as opposed to a gateway that simply presents descriptions of information sources and content. He disputes whether 'cloning' Google is a benefit for the users, as it may not encourage good evaluation habits.

[1] Covey, Denise Troll. *Usae and usability assessment: library practice and concerns.* USA: Council on Library and Information Resources and DLF, 2002.
[2] Joint, Nicholas. Evaluating the quality of library portals. *Library Review* 2005; 54(6): 337-341.

2 Usability testing methods

2.1 User protocols

2.1.1 Aims

Protocols examine what people do, or don't do when conducting certain tasks at the interface. User behaviour, and how they go about the task are recorded in some way. Data collected can reveal navigational problems, and how users visualize the task in front of them, their mental model of the task.

2.1.2 How it works

Protocol studies usually have a list of tasks that the participants are expected to perform – and typically the sessions may be 60 minutes in length (but this will depend on the organisational setting). Each participant is asked to 'think aloud' while doing the task. There are usually facilitators who encourage the participant to keep thinking aloud, prompting as necessary, but not offering specific advice.

The protocol will specify what behaviour is to be recorded. Some human-computer interaction studies will examine body behaviour. In other protocols the emphasis may be on the order in which participants try various menu selections. Software can be used to capture participant keystrokes – which are often performed very quickly and which observers might easily miss. Usually it is necessary to tape (audio or video) the participant, and take notes of the behaviour. In most settings there will be a facilitator (who acts like a focus group facilitator) plus observers who take notes of the behaviour. The facilitator has to assure the participants that it is the software being tested, not their skills. Observers simply observe, and note what participants do.

It is advisable to collect demographic information and some other data on user satisfaction.

2.1.3 Advantages

Although lengthy to do, the number of subjects required may be few. Many cite Nielsen[3] as recommending that around five subjects per demographic group is sufficient (but Nielsen's own writing explains that for a fuller examination more subjects may be required). The tapes are often used simply to annotate the notes made by observers, so save on transcribing time.

The technique provides information that would be hard to obtain by other methods, and gives a rich picture of user behaviour.

2.1.4 Disadvantages

The tasks need to be chosen carefully, and need to be realistic and relevant to the participants. The facilitator needs to be sufficiently skilled to be neutral, but aware of the main design issues so as not to miss important aspects of the user behaviour. Notes are necessary as audiotapes may fail (or be indistinct) and it may be unclear how the users went about the task. Researchers involved in the protocol analysis need to be reasonably skilled in such work, ideally.

The technique needs to be piloted to ensure that the required data can be collected easily and reliably. Participants need to be willing and able to think aloud. The difficulties in getting subjects to participate should not be underestimated – and even if sessions are booked there will be no-shows.

[3] Nielsen, J. Why you only need to test with 5 users, retrieved from http://www.useit.com/alertbox/20000319.html, on 22 Dec. 05

2.1.5 Examples of usage of the technique

User protocols have been used to change the order and presentation of search results, change the placement of links, page layouts, providing more online help, improvements in navigation and search functionality (as listed in the DLF study[4]).

The University of Saskatchewan[5] used a think aloud protocol with five different types of user for a Web site prototype that had a mock-up graphic of the 'ask a librarian reference service'. Participants were asked five questions, including two that were intentionally difficult to prompt the participants to think of using the 'ask a librarian reference service'. Observers recorded the paths taken to find the information and participants were encouraged to think aloud as much as possible. Each question was timed and participants encouraged to move on to the next question after five minutes on a question. Results indicated that terms such as database and interlibrary loan are not well understood by users. Users tended to scan for words that matched their need.

UCLA[6] asked participants (n=10) to find information about library services, how to obtain an interlibrary loan, how to access electronic resources off campus and how to find a wide range of materials for a research project on a specific subject (7 tasks in total). Two facilitators conducted the individual sessions – one acting as an observer.

Northern Illinois University libraries[7] had 62 subjects who each attempted to complete an 11 item task list. Sessions were taped (only the screen video'd and the participant's voice recorded). A moderator facilitated the session, and noted the time the participant started and completed each task, asking participants to move on after spending five minutes on a task. Participants were given 10 dollars for participation. Analysis noted that some tasks presented participations with problems about the route to be taken. Library jargon was another problem.

Carnegie Mellon[8] used think aloud protocols with user representatives (n=9) working with a live prototype of the redesigned site. Tasks were more directed than in other examples (e.g. Find out about archival collections, and what services the University Archives provides. Return to the home page). The tasks were evaluated using a rating scale (based on Nielsen) that considered the frequency with which a problem occurred, how difficult it was for users to overcome, and the persistence of the problem (one-time or frequent occurrence). The researcher only provided help if the participant got completely stuck, and testing for that task halted. Each participant was audio-taped. Findings indicated that users navigate from top to bottom and left to right. Chunking information in sidebar buckets helped. Limiting descriptive text improves the visibility of the kinks or keywords in the bucket enabling quick scanning.

The University of Wisconsin[9] conducted think aloud sessions with participants (n=5) who carried out five navigational tasks, followed up with a questionnaire. User behaviour was video'd and the 'thinking aloud' recorded. The researcher acted as facilitator. Evaluation was by task, identifying interface problems associated with the task, and common themes then identified and solutions prioritised.

---

[4] Covey, Denise Troll, above reference, p. 22

[5] Duncan, Vicky, Fichter, Darlene M. What words and where? Applying usability testing techniques to name a new live reference service. *Journal of the Medical Library Association* 2004; 92(2): 218-225.

[6] Turnbow, Dominique, Kasianovitz, Kris, Snyder, Lise, Gilbert, David, Yamamoto, David. Usability testing for web redesign: a UCLA case study. *OCLC Systems & Services* 2005; 21(3): 226-234.

[7] VandeCreek, Leanne M. Usability analysis of Northern Illinois University Libraries' website: a case study. *OCLC Systems & Services* 2005; 21(3): 181-192.

[8] George, Carole A. Usability testing and design of a library website: an iterative approach. *OCLC Systems & Services* 2005; 21(3): 167-180.

[9] Clark, Jason A. A usability study of the Belgian-American research collection: measuring the functionality of a digital library. *OCLC Systems & Services* 2005; 20(3): 115-127.

Georgia Tech Library[10] recruited 8 participants for a series of tasks (no information on number, or time taken). Redesign of the site was done to suite the user mental models revealed during the testing. The Georgia team emphasise the importance of continuous usability testing as the site and services develop.

Ping Li[11] describes how full cognitive task analysis should be performed (no results given for the evidence-based nursing sites, unfortunately)

2.2 Heuristic evaluation

2.2.1 Aims

Heuristic evaluation focuses on the compliance with principles of good interface design, using accepted principles produced by design experts such as Nielsen.

2.2.2 How it works

Three to five evaluators are recommended – Nielsen[12] suggests that 80% of design problems can be assessed by five evaluators. Evaluators follow a list of tasks that may indicate how to perform the tasks. The findings are presented as a list of observation of the instances when the evaluator identifies that the interface does not obey good design principles.

2.2.3 Advantages

There is no need to recruit subject participants, or consider the user groups. They may be more comprehensive than user protocols as every aspect of the interface can be assessed. They are useful as a way of measuring the number of keystrokes or clicks to reach certain points of the site, and how, for example, disabled users are able to use the site.

2.2.4 Disadvantages

Skilled evaluators tend to be skilled users of an interface and the user problems may not be visible or obvious to them

2.2.5 Examples of usage of the technique

Catherine Ebenezer[13] assessed the following usability metrics from 15 tasks, given to each of seven participants: percentage of tasks completed, number of false starts for each task, longest time take for each task, number of prompts required per task per user, and user satisfaction ratings. Catherine's research is a bit of a hybrid between the talk aloud protocol which is usually conducted on participants who are typical users, and heuristic evaluation which may often by conducted by fairly skilled interface users and testers. In Catherine's case, the testers were typical users, and the problems in conducting the tests (some tape recording is mentioned but it was not always possible to decipher the tapes, no additional observers apart from the person facilitating the session).

[10] King, Heather Jeffcoat, Jannik, Catherine M. Redesigning for usability: information architetuvre and usability testing for Georgia Tech Library's website. *OCLC Systems & Services* 2005; 21(3): 235-243.
[11] Li, Ping. Cognitive task analysis: a cognitive approach to evaluate evidence-based nursing websites. *OCLC Systems & Services* 2005; 21(3): 252-256.
[12] Nielsen, J. Cost of user testing a Website. Jakob Nielsen's alertbox, 3 May 1998. http://www.useit.com/alertbox.9805053.html
[13] Ebenezer, Catherine. Usability evaluation of an NHS website. *Health Information and Libraries Journal* 2003; 20 (3):134-142.

2.3 Paper prototypes and scenarios

2.3.1 Aims

Paper prototypes and scenarios are very similar to think aloud sessions. The main difference is that paper or 'screen prints' are used to help assess likely navigation routes or interpretation of link labels.

2.3.2 How it works

Large post-it notes can be used on a paper prototype of the interface to provide information about menu selections. Participants are asked to imagine performing particular tasks, or assess interpretation of the link labels, for example.

2 3.3 Advantages

Participants are less threatened than by 'the real thing'. Relatively cheap and easier, possibly, to assess what participants are doing as the behaviour is easier to see.

2.3.4 Disadvantages

There is no good reason to believe that what people do on paper matches what they do when faced with a real interface. The paper may slow them down and they may think and reflect more carefully.

2.3.5 Examples of usage of the technique

The University of Michigan[14] study used paper prototypes of the home page, and a variety of lower level pages using the information on content organisation obtained from the card sorting test (see Section 2.4.5). Ten tasks were created, five for one discipline, five for another. The consultant talked to the participants during the testing, leading them through the tasks and follow-up questionnaires for each task. A library staff member acted as observer, taking notes on the time, path taken, problems encountered.

2.4 Card sorting tests

2.4.1 Aims

Card sorting tests assess how users might organise the resources on a site and what they think the link labels imply.

2.4.2 How it works

Participants organise note cards containing service or collection descriptions into stacks of related information, label the stacks and label the individual service and collection descriptions in each stack. The process can also be done in reverse, by asking participants which label they would use to find a particular service or collection, or by asking participants what they think lies underneath each label, then showing them the collection of resources or services under that label and asking them what they would call it. Nielsen[15] recommends 15 users for card sorting tests.

2 4.3 Advantages

This is a good way of sorting out the library vocabulary problems, as users simply do not understand what is meant by some of the labels commonly used on library Web sites.

---

[14] Tolliver, Robert L., Carter, David S., Chapman, Suzanne E., Edwards, Phillip M., Fisher, Jeanie E., Haines, Annette L., Krolikowski, Lana E., Price, Rebecca M. Website redesign and testing with a usability consultant: lessons learned. *OCLC Systems & Services* 2005; 21(3): 156-166.

[15] Nielsen,J. Card sorting: how many users to test. July 19, 2004. http://www.useit.com/alertbox/20040719.html

2.4.4 Disadvantages

This is not the easiest of techniques to analyse systematically.

2.4.5 Examples of usage of the technique

Catherine Ebenezer[16] asked respondents to do a card sorting exercise, with sets of paper slips, one slip for each item on each of the menus, with menu category headings included. Subjects were asked to sort the slips into categories using one of the menu headings included among the slips or devising their own if they wished. Respondents also completed a reverse card sorting exercise in a questionnaire by asking people what they expected to be included in the main menu and sub-menu items, i.e. what the main menu items would include, and what sort of information the links would indicate.

The University of Saskatchewan[17] used several methods to help decide what to call an online reference 'ask a librarian' service. The first simply asked respondents (n=9, first round and n=7 second round) for their views on the wording and reasons for their preferences for possible labels and graphics. Two rounds of tests were used, to try to come up with a clear winner, refining the labels and graphics in the second round.

The UCLA team printed the names of the links on one side of a white card (numbered). The reverse of each card carried a brief definition. Participants (n=40) were given one hour to organize the 76 cards (for 76 links) into any number of set that made sense to them and to name the card sets. SPSS (plus some semantic matching of similar terms) was used to help in the analysis. Findings illustrated that links to information on subject librarians would be needed from various categories, but that most participants expected 'copying and printing' to appear under a Services category.

For the University of Michigan[18] the usability consultant took terms from the top, second, and third level pages of the site in use, library staff added a few more possible options, resulting in 54 cards. Participants (n=6) asked to sort the cards into groups that made sense to them, and then label the groups by writing an appropriate label on the groups.

2.5 Transaction logging and remote monitoring

2.5.1 Aims

The technique is used to study the navigational routes chosen by users.

2.5.2 How it works

Transaction logs, often for particular test exercises, are printed out and analysed, to count the number of links tried, whether the users found the right link, number of links tried.

2.5.3 Advantages

Potentially, this technique is less labour intensive in terms of the staff time required to facilitate usability sessions and observe users.

2.5.4 Disadvantages

Unlike the think aloud protocol, there may be no record of the reasons why users have chosen to navigate by a particular route.

[16] Ebenezer, as above.
[17] Duncan, Vicky, Fichter, Darlene M. as above
[18] Tolliver, Robert L., Carter, David S., et al. as above.

2.5.5 Examples of usage of the technique

The University of South Florida[19] used transaction logs of 22 students asked to navigate to selected resources (within seven clicks), comparing these with the records (manual notes) of students asked to record their link paths. The automated technique was adapted to use of the Apache Web server (although all data from network traffic outside of the library domain was not recorded), with a script to eliminate the graphics calls developed to make the analysis easier. Two groups of 16 students undertook the formal usability test, and this was followed up with questionnaires on the satisfaction with the interface. Comparison of findings suggested that students were quite happy with the interface despite the evidence of long navigational paths and missed answers.

The California State University San Marcos investigated three software tools (Timbuktu, NetMeeting and Camtasia) for use in remote observation studies. There are ethical problems in observing users remotely.

3 Observations on the techniques

User protocols provide detailed information, they don't (thankfully) perhaps require a large number of people, but it is important to have a range of types of user to reveal the main problems. Indications from the way they have been used in the examples noted suggest that good practice requires:

- Careful selection of tasks – they need to be relevant, allow some degree of success to avoid frustration for the participants, but also cover the range of resources, and likely tasks on the site
- Ideally sessions should be video'd (screen only) and participants encourage to think aloud, with observer(s) available to make notes as well to complement the recording. The facilitator can be the observer but combining two roles can be difficult.
- Simpler timing tasks could be managed by fewer researchers, but would not give a depth of detail often required.
- Participants need a reward for participation
- Participants need to be fully informed, give consent should be debriefed
- A supplementary questionnaire or interview helps give demographic data that may help identify problems for particular user groups. It may also help participants vent their opinions on aspects of the site
- Nielsen[20] recommends using simple satisfaction ratings

Heuristic evaluation may tend to focus too much on the 'looks and layout' of the site. It may be more suitable when users are expert users.
Paper prototypes are not reported in the literature very much but may be relatively easy to manage from the research perspective.
Card sorting exercises seem very useful in identifying vocabulary problems and the type of labelling and organisation that some user groups would find useful (e.g. for more personalised interfaces). They do not require as much supervision as the user protocol exercises.

For **more information and guidance,** consult the following resources:
Norlin, Elaina and Winters, C.M. Usability testing for library web sites: a hands-on guide. Chicago: ALA, 2002.
Jakob Nielsen's web site: http://www.useit.com

[19] Allen, Maryellen. A case study of the usability testing of the University of South Florida's virtual library interface design. *Online Information Review* 2002; 26(1): 40-53.
[20] Nielsen, J. Time budgets for usability sessions. September 12, 2005. Retrieved from http://www.useit.com/alertbox/usability_sessions.html on 22 Dec. 05.