

Chapter 3

Methodology

Introduction

The project's main purpose was concerned with the investigation of the use of transaction logs and other methods for the evaluation of online catalogues. This chapter describes the two phases of the methodological approach adopted for the study. In the first phase, field tests of a basic screen logging facility were undertaken which led to the specification of an enhanced logging version and the formulation of test research questions. A set of evaluative experiments were then carried out in the second phase to explore various methods of addressing the different test research questions.

3.1 Evaluation: the user perspective

The critical review presented in the previous chapter highlights the problems of evaluating library catalogues from the user's perspective. Online catalogues could be evaluated in terms of system performance by testing retrieval effectiveness through search outcomes or measures of recall and precision. However our concern here is to focus on the user and the insight which could be gained into user information seeking behaviour, particularly through transaction log analysis as well as other methods of eliciting information from users. User performance is thus seen as the means of evaluating catalogue performance. Moreover a holistic approach to the information retrieval task is also adopted, whereby the entire search process is taken into account, including activities which lead to or follow the catalogue consultation.

3.2 Transaction log analysis limitations

The methodological shortcomings of transaction log analysis were touched upon in section 2.2.4. These include difficulties in identifying individual search sessions, the sheer volume of data to process and analyse and the limitations in the type of data collected. Results to date have provided little qualitative insight on how the data can be interpreted in terms of user searching behaviour and the interactive nature of the information retrieval task.

3.2.1 Demarcation of user sessions

One of the drawbacks in extracting more meaningful data has been the inability to demarcate user sessions on public access terminals. Direct user observation in conjunction with log data has been the main method of determining search boundaries, (Borgman 1983; Mitev, Venner & Walker 1985; Barber & Riccalton 1988). Based on matching log timings with terminal activity, this method nevertheless requires a certain amount of guesswork.

3.2.2 Qualitative analysis of individual searches

In the Okapi'86 project, logs were extensively analysed manually to identify individual searches within sessions, which could then be repeated for retrieval tests on different versions of the prototype catalogue. An attempt was also made to categorize individual searches in relation to each other within a single user session, i.e. as broader, narrower or related. This data however, was not further explored, (Walker & Jones, 1987).

3.2.3 Validating transaction logs with other data gathering methods

An attempt to validate findings from log analysis with other methods of data collection or vice versa, was tentatively made in a study on the LS2000 catalogue at the University of Newcastle, (Barber & Riccalton, 1988). Search types and timings were compared in three approaches including a self-administered questionnaire, a structured interview, and logs. Although the findings were found to be comparable across the three methods, this did not provide any information on the strength of the methods in answering more qualitative research questions.

3.3 Improving transaction log analysis

A major concern of the project was to improve on the effectiveness of logs as a method of evaluation through a combination of approaches including:

- a) Capturing a more complete picture of the activity at the terminal, that is not only the user input but also the system output, so that the interactive nature of the task becomes more evident.
- b) Compressing the data captured to allow for more automatic means of analysis.
- c) Enhancing the basic data collected at the terminal by eliciting more direct information from the user in the course of the search, i.e. by online interceptions on a front end.
- d) Validating transaction logs with other methods of data collection. The purpose was to establish what user information could be obtained from transaction logs and what could be best extracted by other methods.

3.4 Basic screen logging facility

The first phase of the project concentrated on the development of a basic transaction logging facility. Unlike previous practice of logging terminal activity from the central processor and generating coded notations, the approach adopted was to record all of the screen displays intact without any reformatting for output on a front end. The design of the facility is described in full in Chapter 4.

The purpose was two fold:

- to produce a complete and exact record of screen transactions including user input and system output.
- to have the logging facility double as a replay facility whereby searches could be replayed to users.

3.4.1 Field testing of the basic logging facility

The testing of the logging facility was undertaken in the Library at City University on the CLSI CL-CAT online catalogue. A PC was installed alongside a bank of OPAC terminals and users who wanted to consult the catalogue were invited to use the PC as an alternative to the other terminals. Although the equipment looked slightly different than the rest of the terminals at hand, users were reassured that they would be searching exactly the same system. Permission was sought to observe them as they interrogated the system and they were also asked to talk aloud as they proceeded. The experimenter initialised the beginning of the user session and marked the end by pressing a function key. A time stamp recorded the length of sessions.

3.4.2 Combined methodology for data collection

An observation and talk aloud technique with real time interpretation on a highly structured dual questionnaire / observation form first used by Hancock (1987), was adapted here for the online environment (Appendix E). This was complemented by the screen logging which provided an exact record of the transactions for printing and analysis. Two samples were collected:

OPAC1 In this sample users were selected as they approached the online catalogue and were invited to participate by searching on the PC as described above. Ninety-one users were asked and no-one declined. Users were also followed to the shelves after the consultation.

OPAC2 In order to include non-OPAC users a second sample was collected. In this sample every third person who entered the library was approached by the experimenter who explained that a survey was being conducted on the use of the library and could they state the purpose of their visit. No mention was made of the computerized library system in order to ensure that library users were not being influenced or in any way predisposed towards the use or the non-use of the catalogue. Those who indicated that they intended to look for library book material in any fashion were invited to

participate. Those who intended to carry out other activities, i.e. photocopy, return books, study or look for periodicals, were simply thanked for stating the purpose of their visit and were not selected. Eight persons who could have qualified, declined to participate whereas one hundred others accepted.

The one hundred who qualified were asked to carry on with their intended searching activity and permission was sought to accompany them and observe as they carried out their task.

The first objective of the field tests was to determine a possible specification for an enhanced logging facility. The data collected from the 191 cases which included 204 individual searches made it possible to:

- identify different types or categories of searches,
- indicate possible points in the search process where the front end could intercept the searcher to elicit appropriate information,
- provide some indication of what information an enhanced transaction log would need to obtain in addition to what is captured by the basic log

The specification of the enhanced transaction logging facility and its functions are discussed in Chapter 4.

The second objective was to ascertain what evaluative test research questions could be formulated in order to carry out different experiments to investigate data gathering methods.

3.5 Test research questions

Results from the initial test of the basic logging facility also served to focus on some of the possible test research questions on the evaluation of online catalogues. The aim was to concentrate on searching behaviour in terms of search formulation and search strategy, thus responding to the conceptual vocabulary problems and retrieval problems encountered by the user as opposed to the mechanical or interface problems.

Because of the interdependence between research questions and methods, the conceptual element of each research questions was coupled with a methodological component. The project had a dual interest in methodology:

- a) to ascertain which method can best provide an answer to a particular evaluative research question,
- b) to combine methods to isolate variables and to enable the validation of evidence in the evaluative process.

In the formulation of research questions, it was recognized that there were different levels or degrees of difficulty in addressing different questions, as well as in assessing the effectiveness of evaluative methods pertaining to specific research questions. There were therefore some inherent limitations as to how thoroughly each test research

question could be tackled. Moreover it was not possible to test all possible appropriate methods for each individual question or ascertain the full potential of each method.

The results of the field testing (Experiment 1) are presented in Chapter 5. Three research questions were formulated from these trials.

1) Subject searching:

What are the characteristics of subject searches and how can these be differentiated from specific item searches?

By what method can subject searches be identified?

Results from the trials revealed that one third of searches initiated as a specific item search at the catalogue were or developed into subject searches. We were also interested in defining whether a searcher who works through a reading list for a number of specific items on a single topic is in fact doing a subject search.

2) Search formulation:

What is the relationship between a searcher's query formulation in interrogating the online catalogue and his expressed topic?

By what method(s) can we best discover the user's information need?

Markey (1984) and Hancock (1987) demonstrated that searchers in traditional library catalogues used a broader search formulation to access the catalogue than their expressed topics. Our initial tests seem to differ with a predominance of exact search formulations, i.e. the access terms to the online catalogue being the same as users' expressed topic. There are indications that the user's information need may differ from what is being expressed initially or from the terms used to access the catalogue. If a searcher were given an opportunity to be more discursive about his information need in searching the online catalogue, we may be in a better position to assess the effectiveness of the different search formulation strategies.

3) Search outcomes and success/failure rates:

What is the relationship between success/failure at the catalogue and at the shelves?

By what method can we find out how searchers make relevance judgments?

By following catalogue users to the shelves in the trial test, it was possible to determine whether relevance judgements made at the catalogue resulted in the retrieval of documents on the shelves. Both Markey (1984) and Hancock (1987) in examining subject searching in traditional catalogues, found that the most common strategy was to extract a class number from the catalogue and proceed to the shelves without a specific item in mind. By contrast the preliminary trials showed that users followed up at the shelves, specific references found in the online catalogue. It would thus appear that more relevance judgements are being made at the online catalogue.

3.6 Preliminary and evaluative experiments

The test research questions derived from the field test were addressed through a set of different experiments. A full account and results of the preliminary trial experiment and the subsequent evaluative experiments are presented in Chapter 5. A summary of the methodology of each is given below.

3.6.1 Experiment 1: holistic approach, observation and talk-aloud

In an attempt to gain a holistic view of the searching activity, the preliminary trial experiment combined the use of the basic screen logging facility with a talk-aloud technique to observe library users at the online catalogue and at the shelves.

3.6.2 Experiment 2: offline / online, pre-search / post-search questionnaires for subject searching strategies

In the second experiment pre-search and post-search, online and offline questionnaires were used together with the screen logging facility to examine searchers' strategies in undertaking a subject searching. The object was to validate the information obtained from the logs with that elicited from users by the different questionnaires.

3.6.3 Experiment 3: offline / online, pre-search / post-search questionnaires for search articulation and formulation

The third experiment aimed to encourage users to express their search requirements as fully as possible. From the paired combinations of questionnaires as in Experiment 2, and the transaction logs, the written topics declared by the searchers were compared with the search terms used to access the catalogue.

3.6.4 Experiment 4: OLIVE interactive questionnaires

Whereas the previous two experiments explored the reliability of questionnaires prior to or after a search, Experiment 4 considered the effect of introducing similar questionnaires during the course of a search. The enhanced logging facility allowed for the front end to intercept and for certain questions to be put to the searcher if certain conditions were met during the search process. Although the two questionnaires which were tested replicated that of the second and third experiments for subject searching strategies and search articulation, the interactive version targeted specific outcomes within search sequences.

3.6.5 Search replays

The fifth experiment investigated the use of the search replay facility. Searchers were invited to comment on their search as it was being replayed to them in real time. The

purpose was to attempt to establish how they viewed their search terms in relation to their expressed topics, on what basis they made relevance judgments on the references displayed and to what extent they were satisfied by the search outcome.