

Chapter 6

Assessment of Evaluative Methods

Introduction

Whilst taking into account the results of the experiments reported in the previous chapter, the following discussion primarily aims to provide a critical assessment of the evaluative methods applied to address the different research questions. In testing the various methods and approaches, the criteria which were considered included, a) the information which could be elicited from the user, b) the validation of the data by other methods, and c) the practical limitations in applying the methods.

6.1 Holistic vs ‘black-box’ approach

The holistic approach adopted in Experiment 1, which included the observation of searching at the library shelves by catalogue users and non-users, revealed a) a predominance of subject searching, b) an inherent need for subject searching and c) the limitations of the library catalogue in meeting that need. This brings into question the dual assumption of information need and the use of the library catalogue in meeting that need. The ‘black-box’ approach of previous studies emphasized the distinction between known-item and subject searching at the library catalogue, and concluded that the use of the catalogue as a finding tool reflected users’ basic requirement for specific item searching, thus minimizing the need for subject searching.

The study of information seeking in the context of the library environment as a whole also brought to light the role of known citations in subject searching. The identification of hybrid searches whereby one in three searches initiated as specific item searches developed into subject searches either at the catalogue or at the shelves is clearly significant (Section 5.1.2).

Although following users to the shelves and directly observing the searching activity at the shelves proved to be fruitful, it is an approach which is very time consuming and can only realistically be applied in a limited way. Hence the project also considered other methods to see which could be applied more easily and also be as informative.

6.2 Observation combined with talk-aloud technique

By adopting a holistic approach it was possible to record observable behaviour in a continuous fashion from the beginning to the end of a search. The talk-aloud technique applied in Experiment 1, was used to complement or confirm what was being observed without having the user provide any explanations for the actions taken. Searchers were asked to talk-aloud but were left to choose their own level of verbal comment so that the amount of information volunteered varied with individuals and with different search situations.

Great care was taken by the experimenter not to interrupt but to leave the searcher in control. As such there appeared to be no evidence of interfering with the continuity of the search process. Whether or not interference could be manifested in other ways has not been determined. In leaving the level of interaction between the experimenter and searcher at a minimal level, the method did not inform on searchers' full intentions, motivation or perception of the task or problems encountered. It may have been possible to have gathered the same information by observation alone particularly at the shelves, but being observed in silence may have been more intrusive. Equally it may have been more disruptive for the searcher if the experimenter had interrupted intermittently.

At the online catalogue, the verbal data provided by the user was essential to indicate choices and decisions made in the course of the consultation and particularly the information extracted for further searching at the shelves. Through the combined data gathering methods of observation and talk-aloud, the object was to record the full progression of searches. Any attempt to seek more indepth information, relating for example to search formulations, would certainly have encroached on the searching process.

Searchers did not have any difficulty in providing a running commentary. At the catalogue the experimenter recorded the data which would assist in observing the searcher at the shelves, thus duplicating some of the data already recorded in the log. The real time interpretation and recording of data on a structured form served as an efficient way of getting around the tape recording, transcription and coding associated with the use of protocol analysis.

6.3 Basic screen logging facility

The screen logging facility used in Experiment 1 provided a comprehensive source of data of the searching activity at the terminal but did not include any data pertaining to the user's interpretations, judgements or intentions. The latter was obtained from the user's verbal data. Hence the analysis of each search required both sources of data.

Interpreting the logs independently from the recorded verbal data even with user sessions being clearly demarcated still proved to be difficult, particularly in deciphering individual searches within a search session and establishing whether a session pertained to an individual search or multiple search. In fact there were very few multiple search sessions. In the vast majority of cases what in the log may appear as possible separate

searches, were in fact interrelated.

6.4 Combining screen logging with questionnaires

Experiments 2 and 3 explored the possibility of complementing screen logs with information obtained from written questionnaires as a possible alternative to the holistic, observation, talk- aloud approach adopted in Experiment 1. The questions were thus designed to inform not only on the catalogue search but on the information seeking activity as a whole.

In addressing the question of subject searching strategies, the results in Experiment 2 show users' clear intent to undertake subject searches via a specific item search, i.e. a hybrid search, irrespective of when (pre-search or post-search) or how (offline or online) the questionnaire was put to them. These findings which are comparable to the results of Experiment 1, confirm in turn the reliability of eliciting this type of information from users' through the combined methods of questionnaires and screen logs. Users' intent appears to be independent of the catalogue consultation, i.e. users expect to undertake subject searching at the shelves in addition or irrespective of the catalogue consultation.

However, if the logs are taken on their own, just over a third of the searches do not in fact represent users' full intent. It would appear that basic screen logs alone are not a reliable source for identifying types of searches and are thus not an adequate tool for monitoring searching behaviour.

The strength of the combined methods of eliciting information from users from questionnaires as well as the logs is also confirmed in Experiment 3, where users were asked to articulate their information need in a formalised statement. Again there seemed to be little difference between the information elicited before or after the catalogue consultation online or offline.

However seeking that type of information from the user was difficult. Although 48% of searchers did distinguish between their search articulation and search formulation to access the catalogue, 52% did not. Encouraging users to articulate a 'true' or full representation of their need (if known) remains problematic. It would seem that this type of information is too complex to elicit by using a direct open ended written question. Nevertheless both of these experiments established the reliability of pre-search and post-search online questionnaires.

6.5 Online interactive questionnaires (Olive)

In Experiment 4, the online interactive version of the pre-search and post-search questionnaires used in Experiments 2 and 3, tested the viability of intercepting the user in the course of a search. Although the samples are small, the results do not seem to differ from those obtained from the pre-search and post-search questions. There were no obvious indications that the interceptions and question interfered with the searching process.

Interceptions could only be triggered after a restart, and conditions for the questions were limited to events in the previous chain or search sequence. It soon became apparent that the combination of determining other appropriate interception points in addition to setting conditions across chains, would prove to be very complex in practice and would require extensive experimentation. Moreover the implementation of multiple interceptions and multiple sets of conditions would also necessitate a set of clearly defined research questions as well as questionnaires carefully designed in putting appropriate and well formatted questions to the user.

There is scope in further testing the existing facility more vigorously. Although the development of more complex mechanism would seem to be feasible this should be paralleled with prerequisite information on users' internal cognitions.

It must also be borne in mind that the necessary testing of such a facility in an operational library setting to take into account a sufficient number of cases for each of the experiments particularly for paired sets, requires an considerable amount of effort and resources. The five experiments conducted in the present project included around six hundred individual cases.

6.6 Search replays

Although the use of replay facility in Experiment 5 was exploratory, this method demonstrated much potential and deserves further investigation.

Users had no difficulty in recalling the motivations underlying their search as they saw it replayed. At the same time users did not perceive the replay as a new search and actively search the second time around. The method clearly provided greater insight into user's perception of the searching process than any of the other methods described above.

Through the replay facility it should be possible to focus on specific aspects of the user interaction. This in turn could contribute to the formulation of more appropriate questions for use on a wider scale in online interactive questionnaires.