This is What I'm Doing and Why: Reflections on a Think-Aloud Study of Digital Library Users' Information Behaviour

Stephann Makri
Department of Information Studies
University College London
Gower Street, London WC1E 6BT
S.Makri@ucl.ac.uk

Ann Blandford and Anna L. Cox
UCL Interaction Centre
MPEB 8th Floor, University College London
Gower Street, London WC1E 6BT
{A.Blandford, Anna.Cox}@ucl.ac.uk

ABSTRACT
Many user-centred studies of digital libraries include a think-aloud element – where users are asked to verbalise their thoughts, interface actions and sometimes their feelings whilst using digital libraries to help them complete one or more information tasks. These studies are usually conducted with the purpose of identifying usability issues related to the system(s) used or understanding aspects of users' information behaviour. However, few of these studies present detailed accounts of how their think-aloud data was collected and analysed or provide detailed reflection on their methodologies. In this paper, we discuss and reflect on the decisions made when planning and conducting a think-aloud study of lawyers' interactive information behaviour. Our discussion is framed by Blandford et al.'s PRET A Rapporter ('ready to report') framework – a framework that can be used to plan, conduct and describe user-centred studies of digital library use from an information work perspective.

Categories and Subject Descriptors
H.5.2 [**User Interfaces**]: *Evaluation/methodology.*

General Terms
Human Factors

Keywords
Think-aloud, methodology, user study, reflection

1. INTRODUCTION
Several user-centred studies of digital libraries have adopted variants of the think-aloud technique - where participants have been asked to verbalise their thoughts, interface actions and sometimes their feelings when using one or more digital libraries to help them complete an information task or tasks. Some of these studies were conducted with the aim of identifying usability issues associated with the digital libraries used (e.g. [4, 31]). Other think-aloud studies have had the aim of identifying and understanding users' information behaviour or specific aspects of it. However, Nielsen et al. [27] highlight that most papers that present think-aloud studies *"do not discuss in detail what they did, nor reflect on the technique"* (p. 102). This view is shared by Hoppmann [16], who argues that researchers should pay *"more attention to detail"* (p. 211) when describing the design, method and analysis of qualitative think-aloud studies of electronic information environments. This not only suggests the need for thorough discussion of the methodology employed in think-aloud studies of information behaviour, but also the need for *reflection* on methodological decisions made in order to assist researchers in planning and conducting these types of studies in the future.

In this paper, we discuss and reflect on the methodology employed in a study of 22 practicing lawyers working in the London office of a multinational law firm, who were asked to think-aloud whilst using one or more digital libraries to 'find information currently or recently needed for [their] work.' We begin by reviewing existing think-aloud studies of digital libraries with an aim of identifying and understanding information behaviour. This is followed by a discussion and reflection on our methodology, which is framed around Blandford et al.'s 'PRET A Rapporter' ('ready to report') framework [3] – a framework that can be used to plan, conduct and describe user-centred studies of digital library use from an information work perspective.

2. EXISTING THINK-ALOUD STUDIES OF INFORMATION BEHAVIOUR
There are many existing think-aloud studies of digital library users' information behaviour, conducted in different ways to address different research purposes. Some of these studies aimed to describe or model users' broad information behaviour (e.g. [18, 32]). Others sought to investigate particular aspects of information-seeking and use, such as digital library users' search behaviour ([15, 17, 25]), their 'keeping' and 're-finding' behaviour [9] or their relevance selection behaviour [15]. Some of these studies also sought to examine information behaviour from different perspectives; such a cognitive and affective perspective [25] or a mental models perspective [23]. Other information behaviour-focused think-aloud studies sought to examine the impact of particular study-related factors such as an evolving search task [30] or interface used [24] on users' interactive behaviour. There have also been think-aloud studies involving digital library use with a purpose of investigating aspects of the

ACM/IEEE Joint Conference on Digital Libraries'10, June 21–25, 2010, Gold Coast, Queensland, Australia.

think-aloud procedure itself. These studies [e.g. 7, 8] have typically resulted in the identification of information behaviour as a by-product.

Indeed, Branch's study of adolescents using an electronic encyclopaedia [7], which compared think-alouds and think-afters, has been described as 'exceptional' [27, p. 104] due to the high level of detail in which she describes the data collection and analysis process. For example, regarding data collection, Branch discusses the number of participants recruited (3 boys and 2 girls) and their demographics and background (they were aged 12-15, from academic families with computers and electronic encyclopaedias at home). Branch also describes the location and room setup (a lecture room at the University of Alberta equipped with a laptop, tape recorder and two chairs in front of the computer), along with the session length (around 45 minutes) and privacy and confidentiality issues, such as providing the participants with the opportunity to withdraw from the study at any time. Branch also provides details of the think-aloud and think-after procedure employed – such as providing participants with the opportunity to ask questions before the study and to practice thinking aloud and using the system. Details are also provided of the search tasks the adolescents were asked to perform (e.g. 'describe the male cardinal bird' and 'find in what year Queen Elizabeth II was born'). Branch also discusses the nature and amount of researcher intervention (the researcher did not intervene unless the participant asked her a question).

Regarding data analysis, Branch describes how she coded the think-aloud data; she explains that the codes emerged by her reading the think-aloud transcripts as the data collection progressed and after it was complete, clustering major ideas, unique concepts and leftover categories. She then 'cut and paste' parts of the transcripts and grouped them by code. In order to determine the amount of data generated by the think-aloud and think-afters, she also counted the number of words verbalised by participants related to each search task.

Some other think-aloud studies of digital library users' information behaviour also discuss their methodologies in reasonable detail. For example, Hirsh [15] asked 10 elementary school children to think aloud whilst using their choice of an online catalogue, an electronic encyclopaedia, an electronic magazine index, the Yahooligans search engine and a selection of other Internet resources to find information related to an on-going class assignment on sports figures. The purpose of this study was to explore the search strategies they employed and the relevance criteria they adopted. The children were asked to think aloud whilst collecting information for their project. During the think-aloud, they were asked open-ended questions to probe how they were making their relevance decisions. Some of the questions included 'what are you doing now?,' 'now what are you thinking?' and 'why did you try that title?' The researcher did not provide any assistance to participants during the task and shadowed them if they decided to browse the bookshelves for paper-based information. The researcher also took notes on the children's non-verbal behaviour to supplement the audio transcripts.

Manglano, Beaulieu et al. [24] also provide some useful detail on the methodology employed in their study of medical students and professionals' search behaviour. The medics were asked to think aloud whilst using one of two different interfaces of the medical digital library Medline to help them undertake a self-chosen research task. The aim of the study was to investigate the effect of interface design on the participants' search behaviour. The authors provided the medics with information about the study and confidentiality issues and asked them pre-search questions on their status, training level, experience with computers and electronic databases, search purpose and expectations about Medline's content. Participants were then asked to describe their information need and, whilst thinking aloud, describe their interface actions and what they thought Medline was doing. This was with the aim of identifying misconceptions in the medics' mental models. The session, which was not time-restricted, was concluded with post-search questions on their understanding of aspects of the interface, their satisfaction with the interaction and any search difficulties they encountered. The authors' initial coding scheme was theory-driven – based on Fidel's 'search moves' [13] and Bates' 'search tactics' [1]. However, some codes were also data-driven. The codes and their descriptions were iteratively refined to ensure they accurately described the data.

Whilst these studies can be regarded as rare examples of detailed reporting, they only demonstrate a limited amount of reflection on the methodological decisions made when planning and conducting the study. This is also the case with the other studies cited in this section and highlights the need for researchers to include more reflection in published methodologies of information behaviour-related think-aloud studies and the need for more work with the primary aim of reflecting on methodological decisions made. This was the main motivation for writing this paper.

3. THE PRET A RAPPORTER FRAMEWORK

The PRET A Rapporter framework (PRETAR) [3] can be used to both discuss and reflect on a broad range of methodological decisions made when planning and conducting user-centred studies of interactive systems. Whilst PRETAR is not tailored to an information behaviour context (it is intended to be used to plan and describe user-centred studies in general), the authors illustrate through the discussion of several case studies that the framework can be used to describe studies of digital libraries and, more specifically, studies of information behaviour. When describing or reflecting on a user-centred study, the framework involves discussing:

1. The **purpose** of the study – the goals the study sought to address or questions the study sought to answer.
2. The **resources** available for conducting the study and the **constraints** which the study had to work within.
3. The **ethical issues** raised by the study.
4. The **techniques** adopted for collecting data.
5. The **analysis** of the data.
6. How the study was or will be **reported**.

We now discuss and reflect on these considerations in relation to our think-aloud study of practicing lawyers' information behaviour. We cover data collection and analysis decisions in more detail than the other PRETAR stages as the issues raised provide the most

opportunity for reflection. We also, at times, make reference to excerpts from the lawyers' think-aloud transcripts (where 'D' denotes a lawyer working in the Dispute Resolution department and 'T' a lawyer working in the Tax department).

4. DISCUSSION AND REFLECTION ON OUR THINK-ALOUD STUDY

4.1 Purpose of the study

The purpose of our study was to gain a detailed understanding of the interactive information behaviour displayed by practicing lawyers when using digital libraries as part of their everyday work. Our motivation was user centred; we believed that in order to ensure that digital libraries truly support their users, it would be necessary to gain a detailed understanding of their interactive behaviour when using digital library systems to satisfy real information needs. As we sought to gain as realistic an insight of their behaviour as possible, we decided to set the relatively broad information task of 'finding information you currently need or recently needed for your work.' As explained by Blandford et al. [4], who had a similar motivation but sought to identify usability issues related to digital libraries rather than information behaviour, setting a broad information task that allows participants to conduct their everyday work avoids the need for artificial tasks, "*which are liable to be either too precisely defined to be natural or too meaningless for participants" (p. 181).*

Although the broad information task set was highly-related to our study's purpose and resulted in the display of a wide range of information behaviours, the fact that the task demanded that the lawyers '*find* information' served to constrain the behaviours displayed somewhat. Whilst the task did not directly imply active information-*seeking* (as opposed to more passive forms of information encountering such as receiving e-mail alerts), this was implicit in the wording of the question. Similarly, whilst this task was not intended to exclude information use behaviour, it primarily encouraged the display of information-*seeking* behaviour, without much demonstration of how the information found was used as part of their work. We attempted to re-address the balance during data collection by asking wrap-up questions aimed at probing the boundaries of the identified information behaviours and identifying behaviours that were not currently supported by digital law libraries. We discuss the wrap-up questions further in section 4.2.1.2.

Whilst we considered alternative task wordings such as 'show me how you came across information you have used for your work' and 'find and/or make use of information you currently need or have recently needed for your work,' we decided these wordings also had their own inherent problems; we felt the former was too vague and the latter too specific (in the sense that it implied the need for the lawyers to think of a task with a clear feed-in from information-seeking to information use). Therefore both were likely to cause confusion about what the task actually demanded. This was one of several methodological decisions made that involved trading off several potential approaches and making a final judgement based on which approach would, in our opinion, minimise the opportunity for data bias whilst maximising the opportunity for the collection of rich data that would give rise to the identification of a wide range of information behaviours.

4.2 Study resources and constraints

4.2.1.1 Participant recruitment process

We recruited an evolving theoretical sample of 22 practicing lawyers, where the sample size was only finalised during the course of the study. The main consideration for finalising the sample size was that it should be large enough to enable us to gain both a broad and detailed understanding of the lawyers' information behaviour. As the firm was large, we were fortunate that we would be unlikely to run out of potential participants. However, it was necessary to demonstrate sensitivity to the time pressures faced by practicing lawyers – observations often needed to be rescheduled, sometimes on more than one occasion and often at short notice. By being understanding and flexible when arranging observations, we were able to ensure a high level of participation which, in turn, provided us with confidence in the generalisability of our findings.

The lawyers that agreed to take part in our study were recruited from the mainly contentious 'Dispute Resolution' department (where they worked on cases with multiple parties and a dispute to litigate or resolve) and the mainly non-contentious Tax department (where they worked on cases involving one or more corporations but no 'dispute' as such) of the London office of a multinational law firm. We decided to recruit from these two departments as we were advised by a contact in the firm that although both departments made regular use of digital libraries, their information needs (and therefore their digital library usage and resulting information behaviours) were likely to be very different. We found that although differences in information needs certainly existed (for example the Dispute Resolution lawyers were heavily reliant on a broad range of legal cases and legislation and the Tax lawyers more reliant on specialist tax-related legislation and articles), there was much overlap in the information behaviours displayed. We do not, however, regard the recruitment across department as an unnecessary complication to our methodology. Instead, we regard it as a useful indicator (but not firm evidence) that information behaviour might be similar across contentious and non-contentious departments. Hence our recruitment across departments provided us with added confidence in the generalisability of the information behaviours identified across all departments in the London office.

Our sample included both Trainees and Associates where they deemed that digital library use was 'at least sometimes an important part' of their work. No Partners were recruited as time pressures made it difficult for them to commit to taking part (and after some e-mail exchanges with Partners, it became clear that Partners often delegate their information work to Associates or Trainees). We recruited at all levels of the company hierarchy below the level of Partner as we wanted to observe as broad a range of information behaviours as possible. As with our decision to recruit across departments, the decision to recruit across the company hierarchy did not result in considerably different information behaviour or a noticeable increase in 'information expertise' from Trainee to Associate level. However, once again, we do not regret the decision to sample across the hierarchy as it provided us with extra confidence in the generalisability of our findings.

Personal contact with a senior Partner in the firm was invaluable for enabling us to deal with the bureaucratic aspects of setting up the study, such as establishing a non-disclosure agreement and

procedures for contacting participants and feeding back our findings to the firm. A list of Trainees and Associates in each department was provided by the firm and a designated contact was appointed to pre-authorise contact with individual participants, in order to avoid us contacting participants who had particularly high workloads. A personalised e-mail explaining the purpose of the study, the information task that would be undertaken and the duration of the study (no longer than an hour) was sent to each pre-authorised participant. The e-mail also informed participants that findings from the study would be used to inform the design of digital law libraries and would be shared with the firm itself. Several lawyers commented that they were happy to take part in a study that (a) only involved them doing research they were going to do anyway and (b) would hopefully lead to the design or improvement of the digital libraries they regularly made use of (and often found difficult to use). Contacts within the firm proved to be particularly useful in encouraging participation; some lawyers suggested colleagues that might be interested in participating. In addition, once participation in the Tax department, which was smaller than the Dispute Resolution department seemed to 'dry up,' an Associate offered to forward our e-mail to her Trainees, who made time to participate as the Associate had suggested in her e-mail that the study was worthwhile.

4.2.1.2 Setting and equipment

Tax lawyers performed the broad information task at their desks, using their own computers, whilst Dispute Resolution lawyers used a computer in an office set up within their department. Whilst this decision was made to minimise disruption (as Dispute Resolution lawyers often shared their offices), this also prevented access to their personal bookmarks. Whilst this might have had a minor effect on the information behaviours displayed, we decided that minimising disruption was more important that providing access to an own computer for all lawyers – particularly since all computers had access to the same set of digital libraries and other electronic tools. We believe this was a good decision as few of the Tax lawyers used personal bookmarks, even though they had access to them. All digital libraries could be accessed in the normal way, with the exception of one – LexisNexis Butterworths. This was because on many of the Tax lawyers' own computers, the digital library was set to remember their username and password. This setting-related difference served to highlight difficulties in accessing digital libraries (which we found to be an important information behaviour for all of the lawyers). However, in order to avoid this particular access issue preventing Dispute Resolution lawyers from using LexisNexis Butterworths to undertake their chosen information task, those who encountered password difficulties were offered assistance to log in.

As we wanted to observe as broad and realistic a range of information behaviour as possible, we did not constrain the study by focusing on particular aspects of lawyers' information behaviour or by specifying the use of particular digital libraries. Although the firm subscribed to a wide range of digital libraries, the information behaviour displayed by the lawyers was constrained by the functionality offered by the libraries used. We tried to mitigate this issue by asking wrap-up questions after the think-aloud task that probed the nature and boundaries of the identified behaviours and sought to identify behaviours not currently supported by digital law libraries, for example 'what would you do now that you have printed all the cases you thought would be useful?' and 'you mentioned you would read through the Act and make notes. How exactly would you do that?' Asking wrap-up questions, plus the fact that the study allowed a broad range of information tasks to be undertaken and a wide range of digital libraries and library functionality to be used, allowed us to remain confident that we had identified a broad range of information behaviours. However, it also suggested that in order to identify a wider range of behaviours, it might be useful in future to conduct a complementary study (such as a Contextual Inquiry [2]) designed to gain a more comprehensive understanding of lawyers' informal and paper-based information behaviour. So why did we decide only to focus on lawyers' *electronic* information behaviour in the first place? This was mainly for practical reasons; not only did we believe that practicing lawyers would have been unlikely to have had sufficient free time to take part in an in-depth observation, but we were also aware that the law firm had a strongly ingrained cultural practice of protecting client confidentiality. Therefore we believed that, even with a non-disclosure agreement in place, it would have been difficult to obtain agreement for extended observations. We also decided only to focus on electronic information behaviour as we believed that short, focused observations of lawyers attempting to satisfy one particular information need (rather than several different needs, as would be necessary during a longer observation) would minimise the number of observations required in order to meet our aim of collecting a broad, rich set of behavioural data. Whilst we cannot reliably test this hypothesis, the breadth and depth of our findings suggested that we had made a good choice despite the complex trade-offs we were forced to make.

As we did not want to install any of our own software on the firm's network, we decided only to audio (rather than both audio and screen-capture) the lawyers thinking-aloud whilst performing their task. In order that we could accurately recall the lawyers' interface actions when reviewing their think-aloud transcripts, we made time-stamped notes during the study – writing down actions such as 'clicks browser back button' and 'edits search terms to read 'corporation tax dividends.' Most of the time, the audio and notes were sufficient for understanding their behaviour. When this was not the case, we found it useful to mirror users' interactions ourselves on the digital library they used, whilst listening to the recorded think-aloud session and referring to the relevant notes.

4.3 Ethical issues

Blandford et al. [3] highlight that it is good practice to consider issues surrounding keeping participant data as anonymous as possible and respecting participants' confidentiality and privacy. They also highlight that it is good practice to inform participants of the purpose of the study and what will be done with the data. On our informed consent form, we explained the purpose of the study and that it would be audio recorded. We also explained that the study had gained ethical approval from our university ethical committee. We highlighted that the transcriptions resulting from their think-aloud session would be anonymised from the outset. When asked what this entailed, we explained to the lawyers that this involved the censoring of details that could be used to directly identify individuals (particularly names of staff or clients) or the firm (such as precise place names and the name of the firm's in-house knowledge-management database). Whilst complete

anonymisation is unlikely to be possible in studies such as this, where it is necessary to elicit detailed information surrounding the context of an information task, we believe that identifying a particular individual from the transcripts would require considerable effort. We also believe that adopting this procedure was preferable to the alternative of dissuading the lawyers from discussing specific details about their information task and excluding access to particular digital libraries that might reveal the firm's identity. We also highlighted on our informed consent form that the lawyers would be free to review or edit the transcript arising from their study, or request to withdraw at any time (whereby their audio and transcript would be deleted) and that our findings might be used in academic and non-academic publications and presentations. None of the lawyers asked to review their transcript or withdraw.

As maintaining the confidence of the firm (as well and their privacy and confidentiality) was an important concern, we asked the firm to designate a named individual to review all work arising from the study, including this paper. We also highlighted on our informed consent form that data arising from the study would be shared with the firm itself and that the transcripts would be stored in accordance with the UK Data Protection Act 1998 which, in practice, involved safeguarding the data (by storing personal data - i.e. the lawyers' names) separately from their transcripts and deleting this personal information when no longer needed. Whilst the Data Protection Act only covers *personal* information (and therefore not the anonymised transcripts themselves), we also decided to store hardcopies of the anonymised transcripts in a locked cabinet and softcopies on a password-protected computer, as this was in keeping with the spirit of the Act.

Blandford et al. [3] also highlight that *"while immediate respect of individuals is reasonably obvious, less obvious is the need to continue to respect participants' privacy in future presentations of the work and to show similar respect to groups and organisations"* (p. 11). We found that our strict ethical procedures helped us to respect the long-term privacy of the firm and individual participants. We also found that sharing these procedures with the firm and obtaining their agreement to them *before* data collection began helped us to maintain the firm's confidence.

4.4 Data collection technique

4.4.1.1 Why ask participants to think aloud?

As highlighted by Jakob Nielsen in his guide to usability testing, *"the strength of the thinking-aloud method is to show what the users are doing and why they are doing it while they are doing it in order to avoid later rationalizations"* ([26, p. 196]). We decided on the need to elicit verbal data as we believed that this would provide us with as rich and accurate as possible an insight into the lawyers' interactive information behaviour (i.e. *what they were doing* when using digital libraries to complete their task and *why they were doing it*). We believe that employing the think-aloud technique was highly useful for gaining an understanding of lawyers' interactive information behaviour.

4.4.1.2 Think-aloud or think-after?

We also considered carefully whether to screen record the lawyers undertaking their chosen task and ask them to explain their behaviour *after* completing the task (either as an alternative or supplement to asking them to think aloud during the task). Branch's study comparing think-alouds and think-afters [7], discussed in section 2, concluded that whilst think-alouds can provide rich data, some participants may find it difficult to think aloud during tasks that require cognitive processing (such as complex information tasks) and whilst think-afters may be useful in those situations, they *"may be influenced by forgetting and fabrication"* (p. 389). Another trade-off is highlighted by Bowers and Snyder [6], who found that participants who were asked to think aloud concurrently when using an interactive system tended to read what was on the screen and describe the procedures they were following (i.e. describe *what* they were doing) and those who were asked to think aloud retrospectively when using the same system tended to make more reflective statements about *why* they acted the way they did.

Aware that both options were likely to be highly suitable for addressing the purpose of our study and that there were competing benefits and drawbacks associated with each, we decided to ask our lawyers to think aloud *during* the task mostly based on practical reasons. In particular, Ericsson and Simon [12] highlight that eliciting retrospective accounts can be time-consuming. As our study had to be restricted to around an hour (as we did not think lawyers would be able to commit for longer), we did not think this would be enough time to conduct *and* re-play a think-aloud session (especially a session that was long enough to allow the lawyers to attempt their chosen information task in some detail and hopefully result in the demonstration of rich behaviour). Instead, we decided to take the advice of Branch [7], who suggests combining research methods to *"gather the most complete data"* (p. 389). As part of the wrap-up questions asked at the end of the session (which probed the nature and boundaries of the information behaviours identified), we also asked questions that would better help us understand *why* the lawyers performed particular behaviours or interface actions. Questions included 'why did you use a plus sign when conducting your earlier search?' and 'you mentioned using 'masked proxy access' to login to LexisNexis Butterworths. Why did you need to do that?'

4.4.1.3 Instructing participants on completing the information task and thinking aloud

Whilst some books from the Human-Computer Interaction domain (e.g. Dumas & Redish's 'A Practical Guide to Usability Testing' [10]) provide detailed guidance for instructing participants on how to think aloud, an assumption is made that the primary purpose of the think-aloud study is to identify usability issues related to the interactive systems being used, rather than to understand participants' interactive behaviour. As we were unaware of any detailed guidance for conducting think-aloud studies of information behaviour, or interactive behaviour in general, we devised instructions for our study that we thought would best help us achieve our aim of gaining as detailed and accurate an understanding as possible of a broad range of behaviours.

In order for our think-aloud data to be accurate, we needed to ensure that the think-aloud sessions were as *true to life* as possible

(within our study's constraints). To achieve this, after explaining the purpose of our study, we also told the lawyers that it was our aim to observe behaviour that was as natural as possible and that they should undertake their self-chosen information task in the way that they normally would. The lawyers were told that if they chose to step-through a task they had recently undertaken, they should use the task as a springboard - what they did when looking for the information previously was not important. We also reassured them that the study was not a test of their information skills.

After reading and signing the informed consent form, the lawyers were given a few minutes to think of a suitable information task and then asked to describe the context surrounding the task in detail. Whilst often not directly relevant to the lawyers' information behaviour, gaining a detailed understanding of not only the information task, but the motivation behind it was extremely useful for gaining a richer understanding of their information behaviour. After choosing their task, the lawyers were asked to think aloud – verbalising their thoughts and interface actions as they used the digital library or libraries of their choice to undertake the task. The instructions read out to the lawyers are shown in figure 1.

> You will be asked to **find information that you need for your work**. You should undertake a *real* task that you need to use digital libraries to assist you with. If you do not currently need to perform an information task, you can step-through a recently completed task but *do not need to undertake it in exactly the same way as previously*. Your main aim should be to perform your chosen information task in as natural a manner as possible.
>
> You will be asked to **think aloud** whilst undertaking the task (i.e. to verbalise what you are doing as you are doing it and any thoughts going through your head). I will ask some questions during the task about what you are doing and why and some wrap-up questions afterwards. If you have any questions about the observation, feel free to ask them at any time. However, I may not be able to answer certain questions while you're still undertaking the task as I do not want to bias your comments or behaviour.

Figure 1: Think-aloud instructions read out to participants

Although we considered giving the lawyers an opportunity to practice thinking aloud, we thought they might not regard this as a constructive use of their time. Therefore whilst we offered them the option to practice, we did not enforce a practice. Consequently, none of the lawyers expressed a desire to practice. Whilst none of the lawyers demonstrated particular difficulty thinking-aloud, in subsequent think-aloud studies of interactive information behaviour (not yet reported), we have found it useful for the *researcher* to give a short (10-15 second) example of them thinking aloud when conducting a simple Internet search and looking through the results. This minimises the time required to introduce the study, whilst providing a concrete example of how to think aloud when using electronic information environments.

4.4.1.4 Intervening during the think-aloud session

Ericsson and Simon [12] argue that any researcher comment, prompt or question whilst a participant thinks aloud makes their subsequent think-aloud data unreliable as the intervention alters the flow of information in the participant's short-term memory during the task. However, when evaluating the use of interactive systems, Boren and Ramey [5] highlight that it is often necessary to prompt for data about users' expectations or explanations of their interface actions and that sometimes this data, which Ericsson and Simon would deem as 'unreliable,' can be valued over more procedural information. We believe this is also the case when studying users' interactive information behaviour and suggests the need for researchers to weigh up their priorities with regard to deciding whether to stick to Ericsson and Simon's 'no intervention' rule. Indeed, Tamler [29] suggests that in order to decide whether and how much to intervene in a think-aloud session, it is important to examine the *purpose* of the session. He suggests, for example, that if the purpose of the think-aloud session is to collect quantitative usability data, non-intervention may be particularly important. However, *"if the purpose is to gather qualitative date so as to identify significant user interface problems and recommend design solutions, then openly interacting with users in various ways may not only be useful, but also sometimes necessary"* (p. 12). We felt that whilst a traditional think-aloud study with no intervention might provide us with a reasonable insight into *what* the lawyers were doing when using digital libraries to complete their chosen task (i.e. their interactive behaviour), the complicated and cognitive nature of information work meant we would be unlikely to gain a comprehensive and accurate understanding of their behaviour without knowing *why* they performed certain actions. Therefore we decided to ask questions during the think-aloud session where we believed asking the question was (a) likely to provide us with a greater insight into the lawyers' information behaviours and (b) unlikely to bias the lawyers' future actions.

Whilst we asked questions during each of the think-aloud sessions, we took care to only intervene when we thought this would impact positively on our findings; the researcher remained passive for the vast majority of each session. Whilst we cannot be certain that asking questions did not introduce limited action or halo-effect bias, we believe that our interventions resulted in think-aloud data that was far richer than in might have been had we not intervened. It would be interesting to test this hypothesis in a future study.

The questions asked during the think-aloud session took the form of short and seemingly innocuous questions, posed at opportunistic moments during the study - usually to probe the lawyers' interactive behaviour in more detail. At the beginning of the think-aloud session, we also found it necessary to ask questions to probe more detail on the context of the lawyers' chosen information task and the underlying information need, which often gave us a clearer understanding of the goal of the task and the motivation behind it.

During the think-aloud itself, the most common question asked was 'why did you do [x]?' – to gain an understanding of the lawyers' interface actions or to check our understanding or assumptions. For example, when asked why he clicked on a question mark icon beside a greyed-out checkbox on the segmented field search in LexisNexis Butterworths, a Dispute Resolution lawyer explained that he did so to see whether he had access to that particular search feature, since *"quite often there are sections that we haven't subscribed to"* (DR15). We also asked questions to seek elaboration on comments made or interface actions performed. For example, when one lawyer stated that his search had only returned one result, which '[didn't] look relevant' (T1), we asked 'why don't you think the result looks relevant?' The lawyer's answer provides an explanation of his search behaviour:

"Because I actually made a mistake originally. I wanted to refer to the Finance Act 1996 and I forgot to include the '1996,' so only one hit came up, which was a reference for something to do with the 2006 Budget" (T1).

Questions were also asked to seek clarification on a comment made or interface action performed, often with the purpose of checking our understandings and assumptions. For example, when one lawyer spent time reading through the results list in LexisNexis Butterworths, we asked him which part(s) of the screen he was looking at. He told us he was only skimming the result headings, not reading the snippet of text below each heading presenting the search terms in the context of the document. Our intervention did not seem to bias his future actions; he continued to read the headings, and then edited his search. In another instance, when we asked a Tax lawyer whether he decided on the relevance of the search result he had clicked on by 'weighing up' the results in the list and choosing the most promising one, we found out that he was actually performing slightly different behaviour (i.e. making a binary decision about whether to click on each result in turn):

R: Was it the case that you picked the most likely one to be relevant by weighing them up?

T3: No, it was a more gradual step-by-step thing. I looked at the first one and decided it was not relevant, then I looked at the second one and decided it was not relevant. But that was only because there was 4 of them. If there was like 30 I would have probably gaged them all against the other.

Aside from questions aimed at probing the lawyers' interactive behaviour, we also found it useful to intervene when lawyers strayed away from their stated information task, slipped to providing abstract descriptions rather than demonstrating concrete interface actions or forgot to think-aloud. As an example of straying away from the chosen task, one lawyer offered to look at the history of a particular legal case even though this did not seem to be relevant to her chosen task. The researcher politely declined her offer and requested that she continue with her task. A couple of the lawyers also reverted to giving abstract descriptions of their actions, rather than demonstrating those actions. 'Can you show me…' questions were particularly useful for encouraging these lawyers to shift away from abstract descriptions of their behaviour and continue with their task:

T4: Then I did something a bit more general that didn't just talk about Section 12 of the Capital Gains Tax Act, but talked about the remittance basis for Capital Gains Tax in general and I got a couple of more articles from Lexis.

R: Could you show me what you did?

T4: I just browsed through that initial list of 40 results [returns to results list].

Although the lawyers in our study managed to think aloud without difficulty, many of them went quiet when performing highly cognitive activities such as looking through search results or reading through a document. Asking 'what are you doing now?', as suggested by Dumas and Redish [10], was particularly useful for (indirectly) reminding them to resume verbalisation. For example, Tax lawyer T6 read through a section of the Inland Revenue Manual in LexisNexis Butterworths and then browsed through subsequent headings in the manual's contents tree. When asked 'what are you doing now?', he provided useful detail:

T6: I'm just looking in case any of these individual sections really jump out at me as being potentially helpful for the question that I'm considering, which none of them especially do.

Overall, we found carefully-considered intervention to be extremely useful for the purpose of our study – gaining an understanding of lawyers' interactive information behaviour. The resultant think-aloud data was not only rich enough to provide us with a deep insight into the lawyers' behaviour, but also into the motivation behind this behaviour. The data also revealed no specific evidence to suggest that our interventions had biased the lawyers' future comments or interface actions.

4.5 Data analysis

4.5.1.1 Transcribing the think-aloud data

The audio-recorded think-aloud data were transcribed verbatim and anonymised from the outset (i.e. no details that could be used to specifically identify the participant, firm or client were included on the transcript). We found the process of transcribing our own transcripts useful as it helped us gain familiarity with the data. Bold italics were used to denote when a lawyer emphasised a particular statement and square brackets were used to denote pauses of over five seconds. We also found it particularly useful to summarise lawyers' interface level actions (also in square brackets) as this assisted us when trying to understand the behaviour displayed in the transcript. In order to avoid biasing our analysis, we avoided interpreting any of the lawyers' interface actions when summarising them. For example, one lawyer conducted a digital library search for 'undue inflence,' when he intended to search for 'undue *influence*.' Instead of jumping to a quick interpretation of his actions and transcribing them as '[participant misspells the word 'influence,' which leads to an unsuccessful search],' we summarised his actions without interpretation as '[participant enters the search terms 'undue inflence' into the search box and submits the search, but receives no results].'

4.5.1.2 Identifying information behaviour from the think-aloud data

In order to identify information behaviour from our think-aloud data, we employed aspects of Glaser and Strauss' Grounded Theory methodology [14]. Grounded Theory involves systematically gathering and analysing data during the research process [28, p. 12] and is 'grounded' in the sense that the theory is heavily rooted in the data and emerges through the process of cyclic data-gathering and analysis. Therefore Grounded Theory should not be regarded as a data analysis technique per se, but a methodology for data collection *and* analysis – which, according to Glaser and Strauss, should not be regarded as separate processes. This cyclic approach is known as the 'constant comparative method' and is a key tenet of Grounded Theory. Our data collection and analysis approach allowed us to constantly question and revise our understanding of the lawyers' information behaviour during the analysis. We found it particularly useful to use some of the questions we asked during the think-aloud session as a means of checking our evolving understanding (as discussed in the previous section).

After transcribing a particular think-aloud session, we read the transcript sentence by sentence and assigned codes to parts of it that illustrated particular interactive information behaviours. The coding process was achieved by coding parts of the transcripts that appeared to refer to the same type of interactive behaviour with the same label and refining the analysis through a cyclic process of re-reading the transcripts several times, re-naming codes (for example when a better or more precise description of the behaviour could be identified), merging codes (when two identified information behaviours were deemed to actually be the same), splitting codes (when an information behaviour that had previously been coded under one code was deemed to actually be different) and by re-coding parts of the data under a different code name or unlinking data from a particular code (when data no longer appeared to fit the code name that it had been assigned to).

As an example of assigning a code and later re-coding the data, consider the interactive behaviour of looking through the first page of search results – reading either the document titles and/or the snippet below the title and then clicking on the first document that showed potential to be useful. Initially, we coded this behaviour as 'selecting,' which we defined as 'carefully choosing information' (in this case, documents from the results list). As our data collection progressed, however, we noticed that we had coded two distinct types of behaviour 'selecting' – the behaviour described above, where lawyers 'weighed up' the results set (or part of it) and clicked first on the result that showed the most promise and a more general behaviour – where lawyers started from the first result and decided whether or not to click on it, before moving onto the next result and repeating the process. We decided that 'selecting' accurately described *this* behaviour, but that the 'weighing up' results was best described as 'distinguishing' – an information behaviour that Ellis and Haugan define as *"ranking sources or documents according to their relative importance based on own perceptions"* ([11, p. 399]). We therefore looked back through our transcripts and removed the code 'selecting' from parts of the transcript that seemed to demonstrate the 'weighing up' of search results. We then assigned the new code, 'distinguishing,' to these instances. This example serves to illustrate how constantly comparing the interactive information behaviours displayed by different lawyers resulted in a richer theoretical picture and a more accurate description of their behaviour. Indeed, the process of re-reading the transcripts and asking 'what is (really) going on here?' was invaluable for understanding the lawyers' behaviour. An excerpt from a coded think-aloud transcript is shown in figure 2.

T1: On the off chance, I'll further define "Finance Act 1996" [adds '1996' to original search terms of '"Finance Act" AND "schedule 9" and "participator"']. ***(Codes: Search narrowing, Search re-focusing).*** But because there was only one hit and this will make it more narrow, I'd be surprised if anything comes up. [Conducts search and receives no results]. So what I'll do is take out 'participator' and put in the reference to the exact section that I need to find. [Replaces the term 'participator' with "paragraph 2" and re-submits search]. ***(Codes: Search re-focusing)***.

Figure 2: Excerpt from a coded think-aloud transcript

We coded the data using the 'open' and 'axial' coding elements of Grounded Theory in order to identify the interactive information behaviours displayed by the lawyers and how these behaviours might relate to each other. Strauss and Corbin [28] define open coding as *"the analytic process through which concepts are identified and their properties and dimensions are discovered in data"* (p. 101) and axial coding as *"the process of relating categories to their sub-categories, termed 'axial' because coding occurs around the axis of a category, linking categories at the level of properties and dimensions"* (p. 121). However, it is common for researchers employing Grounded Theory to undertake a third stage of coding, 'selective coding' – defined as *"the process of integrating and refining the theory"* (p. 143) and achieved by relating all code categories to a central 'core' category. We made the choice to perform only open and axial coding (effectively 'stopping short' of generating a theory), based on the purpose of our study. Our aim was not to generate theory per se, but to identify a broad range of interactive information behaviours and to understand these behaviours in detail. We did not believe that attempting to identify a 'core' information behaviour and relating the other identified behaviours to it would be useful for this purpose. Indeed, we found it more useful to establish *firm* boundaries between codes in order to definitively categorise particular information behaviours than to establish fluid boundaries by considering the identified behaviours as highly inter-related and each linked to a particular behaviour that was more important to lawyers than the others.

Our decision to undertake only open and axial coding, but otherwise follow the core principles of Grounded Theory raised an important issue. The issue was related to how we describe our methodology and, more specifically, whether we should call it 'Grounded Theory' even though our study did not aim to or end up generating a theory. In this paper, we have taken care not to label our data collection and analysis procedure as a 'Grounded Theory.' Instead, we have tried to be as precise and transparent as possible about exactly what we did and why, relating our methodological decisions back to the purpose of the study.

Although we did not perform integrative selective coding, we found the other aspects of Grounded Theory to be extremely useful. Strauss and Corbin state that "*although researchers may pick and choose among some of the analytic techniques that we offer, the procedures of making comparisons, asking questions, and sampling based on evolving theoretical concepts are essential features of the methodology"* [28, p. 46]. Indeed all of these essential features of Grounded Theory were particularly useful in our study and, we believe, are likely to be useful for other studies aimed at gaining a detailed understanding of information behaviour in general or particular aspects of it.

A final consideration when identifying information behaviour from our think-aloud data was the need to avoid bias during analysis. In particular, we found that the information behaviours identified were similar to that found in many other studies of information behaviour – including studies that had led to the development of theoretical models of information-seeking. Therefore there was the potential to use existing information theory to guide our analysis. However, as Strauss and Corbin assert, *"the researcher does not want to be so steeped in the literature that he or she is constrained or even stifled by it"* [28, p. 49]. Therefore care was taken to avoid simply relating our data to different information-seeking models in order to identify a model or models which fitted

the data best (which might be regarded as 'forcing' as opposed to 'emergence' in Grounded Theory terms).

Instead, the codes we assigned to parts of the think-aloud transcripts were based on our own terminology, and similarities between the types of behaviour described by our codes and existing theoretical models (notably Ellis's behavioural model of information-seeking [11]) emerged from the analysis. This led us to examine our data in the light of Ellis's model, asking questions of our data such as 'are any of the information behaviours we identified amongst lawyers similar to those found by Ellis and his colleagues and, if so, which ones?', 'Which information behaviours identified by Ellis and his colleagues are not present in our data?' and 'which behaviours in our data were not identified by Ellis and his colleagues?' These questions are related to those that Strauss and Corbin suggest should be asked when relating emerging concepts to previous work; *"are these concepts truly emergent, or am I seeing these concepts in the data because I'm so familiar with them? If they are truly emergent and relevant, then how are they the same as and how are they different from, those in the literature"* (p. 50).

To facilitate easy comparison with Ellis's model, we chose to use Ellis's existing code labels when we believed our data reflected identical (or highly similar) behaviour. Our comparison to Ellis's model resulted in the validation of the model in the legal domain and its extension and refinement (see [21]). However, we should stress that we did not seek to do this from the outset; the purpose of our study was to gain a detailed understanding of lawyers' interactive information behaviour in order to inform digital law library design. We regard the validation, extension and refinement of Ellis's model as an important 'theoretical by-product.' Comparing our findings to Ellis's model helped us achieve our purpose by providing us with a useful reference for questioning the data, resulting in what we believe to be a richer understanding of the lawyers' interactive behaviour. Comparing our data to Ellis's findings was also particularly useful as it highlighted a useful level of abstraction at which to describe the interactive behaviour in order to inform design. For example, we noted that digital library developers might feasibly inform design by asking themselves: 'how can we support or better support Ellis's 'distinguishing' behaviour?' We also noted that the same could be said of many of our identified codes (e.g. 'selecting,' defined on the previous page). Ellis's model therefore provided us with useful meta-theory for coding our data. Indeed, this partly inductive and partly deductive stance not only enabled us to relate our findings to previous research but to *actively use existing studies* to help us better understand our data and ways of analysing it.

Overall, we found that using the open and axial coding techniques along with the other core principles behind Grounded Theory was a highly useful way of identifying information behaviour from our think-aloud data. Although we do not seek to downplay the potential value of other qualitative and quantitative analysis techniques, we believe that this methodology and, in particular, the constant comparative method greatly assisted us in gaining a detailed and 'true' an understanding as possible of our data.

4.6 Reporting the findings

Our findings are reported in [20] and fed into the development of two user-centred methods that can be used to evaluate the functionality and usability of digital libraries, which use the information behaviours identified as their theoretical basis. The methods are reported in [22].

We also fed our findings back to the law firm. As the firm was particularly interested in usability issues with their own in-house knowledge management database, we also provided them with anonymised extracts from our transcripts where lawyers used or referred to the firm's database. We also found it particularly useful to feed back our findings in the form of informal presentations to senior members of the firm (and to participants interested in the findings, who were also invited to attend). The findings were also presented as part of a workshop organised with staff working for LexisNexis Butterworths.

5. CONCLUSION

Our think-aloud study of lawyers' interactive information behaviour involved making many difficult methodological decisions and trade-offs. In these situations, we found it most useful to refer back to the *purpose* of our study in order to decide how best to proceed. By having a clear purpose for our think-aloud study, we were able to choose an information task for participants to carry out that was as closely related as possible to our purpose and plan a study that was highly focused on achieving that purpose – making best use of the resources available whilst taking important constraints and ethical issues into account. Referring back to the purpose of our study also assisted us when collecting and analysing our think-aloud data – enabling us to maximise the likelihood of our study resulting in rich, useful think-aloud data whilst minimising the likelihood of data bias.

Our discussion and reflection on the study not only serves as a form of guidance to researchers who are considering planning a think-aloud study of interactive information behaviour, but also to highlight the important need for rigorous discussion and reflection on the methodology employed and methodological decisions made in future information behaviour studies.

6. ACKNOWLEDGMENTS

This work was supported by an EPSRC Studentship and an ESRC Post-Doctoral Research Fellowship.

7. REFERENCES

[1] Bates, M. (1979). Information Search Tactics. Journal of the American Society for Information Science 30(4), pp. 205-214.

[2] Beyer, H. & Holtzblatt, K. (1998). Contextual Design: Defining Customer-centred Systems. Morgan Kauffman. London, UK.

[3] Blandford, A., Adams A., Attfield, S., Buchanan, G., Gow, G., Makri, S., Rimmer, J. & Warwick, C. (2008). The PRET A Rapporter Framework: Evaluating Digital Libraries from the Perspective of Information Work. Information Processing and Management 44(1), pp. 4-21.

[4] Blandford, A. Stelmaszewska, H. & Bryan-Kinns, N. (2001). Use of Multiple Digital Libraries: A Case Study. In Proceedings of JCDL'01, Roanoke, VA., pp. 179-188. ACM.

[5] Boren, T. & Ramey, J. (2000). Thinking Aloud: Reconciling Theory and Practice. IEEE Transactions on Professional Communication 43(3), pp. 261-278.

[6] Bowers, V. & Snyder, H. (1990). Concurrent Versus Retrospective Verbal Protocols for Comparing Window Usability. In proc. 34th Meeting of the Human Factors Society, pp. 1270-1274. Santa Monica, CA. HFES.

[7] Branch, J. (2000). Investigating the information-Seeking Processes of Adolescents: The Value of using Think Alouds and Think Afters. Library & Information Science Research 22(4), pp. 371-392.

[8] Branch, J. (2001). Junior High Students and Think Alouds: Generating Information-Seeking Process Data Using Concurrent Verbal Protocols. Library & Information Science Research 23(2), pp. 107-122.

[9] Bruce, H., Jones, W. & Dumais, S. (2004). Keeping and Re-Finding Information on the Web: What do People do and what do they Need? In proc. 67th ASIST annual meeting. Chicago, IL. Information Today.

[10] Dumas, J. & Redish, J. (1999). A Practical Guide to Usability Testing. Intellect Books. Exeter, UK.

[11] Ellis, D. & Haugan, M. (1997). Modelling the Information-seeking Patterns of Engineers and Research Scientists in an Industrial Environment. Journal of Documentation 53(4), pp. 384-403.

[12] Ericsson, K. & Simon, H. (1993). Protocol Analysis: Verbal Reports as Data (Rev. Ed.). MIT Press. Cambridge, MA.

[13] Fidel, M. (1985). Moves in Online Searching. Online Review 9(1), pp. 61-74.

[14] Glaser, B. & Strauss, A. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago, IL. Aldine Publishers.

[15] Hirsh, S. (1999). Children's Relevance Criteria and Information Seeking on Electronic Resources. Journal of the American Society for Information Science (14), pp. 1265–1283.

[16] Hoppmann, T. (2009). Examining the 'Point of Frustration.' The Think-Aloud Method Applied to Online Search Tasks. Quality and Quantity 43(2), pp. 211-224.

[17] Jenkins, C., Corritore, C. & Widenbeck, S. (2003). Patterns of Information Seeking on the Web: A Qualitative Study of Domain Expertise and Web Expertise. IT & Society 1(3), pp. 64-89.

[18] Jones, Y. (2006). "Just the Facts Ma'am?" A Contextual Approach to the Legal Information Use Environment. In proc. 6th ACM Conference on Designing Interactive Systems, pp. 357-359. University Park, PA.

[19] Komlodi, A., Soergel, D. & Marchionini, G. (2006). Search Histories for User Support in User Interfaces. Journal of the American Society for Information Science and Technology 57(6), pp. 803-807.

[20] Makri, S. (2009). A Study of Lawyers' Information Behaviour Leading to the Development of Two Methods for Evaluating Electronic Resources. Unpublished doctoral thesis, University College London, UK.

[21] Makri, S., Blandford, A. & Cox, A.L. (2008). Investigating the Information-Seeking Behaviour of Academic Lawyers: From Ellis's Model to Design. Information Processing and Management 44(2), pp. 613-634.

[22] Makri, S., Blandford, A. & Cox, A.L. (2008). Using Information Behaviours to Evaluate the Functionality and Usability of Electronic Resources: From Ellis's Model to Resource Evaluation. Journal of the American Society for Information Science and Technology 59(14), pp. 2244-2267.

[23] Makri, S., Blandford, A., Gow, J., Buchanan, G., Rimmer, J. & Warwick, C. (2007). A Library or Just another Information Resource? A Case Study of Users' Mental Models of Traditional and Digital Libraries. Journal of the American Society for Information Science and Technology 58(3), pp. 433-445.

[24] Manglano, V., Beaulieu, M & Robertson, S. (1998). Evaluation of Interfaces for IRS: Modeling End-User Searching Behaviour. In proc. 20th BCS-IRSG Colloquium, pp. 137-146. Grenoble, France.

[25] Nahl, D. & Tenopir, C. (1996). Affective and Cognitive Searching Behavior of Novice End-Users of a Full-Text Database. Journal of the American Society for Information Science and Technology 47(4), pp. 276-286.

[26] Nielsen, J. (1993). Usability Engineering. Academic Press. Boston, MA.

[27] Nielsen, J., Clemmensen, T. & Yssing, C. (2002). Getting Access to What Goes on in People's Heads? – Reflections on the Think-Aloud Technique. In proceedings of 2nd NordiCHI conference, Aarhus, Denmark, pp. 101-110. ACM, NY.

[28] Strauss A. & Corbin J. (1998). Basics of Qualitative Research. Sage Publishers. London, UK.

[29] Tamler, H. (1998). How (Much) to Intervene in a Usability testing Session. Common Ground 8(3), pp. 11-15.

[30] Vakkari, P., Pennanen, M. & Serola, S. (2003). Changes of Search Terms and Tactics While Writing a Research Proposal: A Longitudinal Case Study. Information Processing and Management 39(3), pp. 445-463.

[31] Van den Haak, M., de Jong, M. & Schellens, P. (2004). Employing Think-Aloud Protocols and Constructive Interaction to Test the Usability of Online Library Catalogues: A Methodological Comparison. Interacting with Computers 16, pp. 1153-1170.

[32] Yang, S. (1997). Information Seeking as Problem-Solving: Using a Qualitative Approach to Uncover the Novice Learners' Information-Seeking Processes in a Perseus Hypertext System. Library & Information Science Research 19(1), pp. 71-92.