Book Selection Behavior in the Physical Library: Implications for eBook Collections

Annika Hinze
Computer Science Department
University of Waikato
Hamilton, New Zealand
hinze@cs.waikato.ac.nz

Dana McKay
Library, The Swinburne Institute
for Social Research
Swinburne University of
Technology
Melbourne, Australia
dmckay@swin.edu.au

Nicholas Vanderschantz,
Claire Timpany,
Sally Jo Cunningham
University of Waikato
vtwoz, ctimpany,
sallyjo@cs.waikato.ac.nz

ABSTRACT
Little is known about how readers select books, whether they be print books or ebooks. In this paper we present a study of how people select physical books from academic library shelves. We use the insights gained into book selection behavior to make suggestions for the design of ebook-based digital libraries in order to better facilitate book selection behavior.

Categories and Subject Descriptors
H3.7. Digital Libraries: User issues;
H5.4. Hypermedia: User Issues

Keywords
Digital libraries, ebooks, human-computer interaction, information seeking behavior.

1. INTRODUCTION
Rowlands, *et al.* noted in 2007 that there is surprisingly little data on how readers select books; a data deficit that remains true today [33]. This dearth of data on how readers select books makes it difficult to design interfaces that adequately support book selection, an activity that has been shown to rely heavily on serendipity in physical environments [26, 35]; unfortunately, this strategy simply doesn't work online [27].

Using information seeker behavior in physical environments to inform DL design is a well established technique: [14] uses music shopping behavior to inform DL design; and [13] uses ethnographic observation to determine the limitations of a digital library for technical support workers. Closer to our work, [11] considers DL design with respect to the reference interview, [8] uses observations of book shoppers to further understanding of book seeking, and [31] reflects on e-reader design with respect to print reading behavior. In this vein, we seek to use book selection behavior in libraries to better inform the design of ebook-based DLs.

Getting a library book is typically a four step process:

(1) identifying books of interest (e.g. via the catalogue),
(2) physically retrieving the books from the shelves,
(3) choosing among the available options, and
(4) accessing and reading the book for the desired content.

The first step, searching for books (or indeed any kind of information) is a well understood behavior. There are numerous studies of information search processes (e.g., [17, 22]) and similarly plentiful studies of how users search for books in the catalogue (e.g., [10, 18, 37]). Previous literature also exists on how readers identify interesting fiction books [30], the types of questions they ask librarians [11, 29], and the kinds of queries posed in bookshops [8].

While the fourth step, reading, is the subject of less research, nonetheless that research is relevant here. Some facets of reading, such as the ability to easily annotate or flip through a text [1] may affect the reader's preference of reading medium [31]. There are a number of studies (e.g., [31]) on the use of e-readers, but little work on the physical act of reading books, perhaps because this activity is difficult to study without being 'creepy' [23]. As reading behavior is relatively unlikely to impact book selection (outside of the medium), it will not be further considered here.

The third step, choosing among available resources, is poorly studied for books, particularly in physical environments [33]. While there is considerable literature on document triage for journal articles (see Section 2.3), this literature does not translate well to book selection because it shows readers using article-specific features (e.g., the abstract) to make decisions [2, 3, 9]. Book selection is largely ignored in the literature on scholarly information seeking, though there have been some studies in artificial environments of decision making over both print books [36] and ebooks [19]. Natural behavior, however, is often very different from that observed in experimental settings [15]; for example, book selection literature does not reflect the shelf-based serendipitous discovery academics value in libraries [35]. Finally, the second step, retrieving a book from the shelves, has been overlooked as not being of significance to the book selection process. We argue, however, that at this point the reader is typically verifying whether or not the potentially useful book is actually relevant to their information needs. Further, during physical retrieval of the book, a reader typically also browses the adjoining shelf space for other books of interest and, again, uses physical cues to decide which books to take home. Surprisingly, which cues readers use to determine whether or not books are useful at the shelves has not been formally investigated.

The work reported in this paper seeks to understand what factors contribute to decision making when selecting books by observing readers selecting books in physical libraries and interviewing them about their behavior. Section 2 of this paper outlines related work, Section 3 describes our methodology, and Section 4 gives our results. In Section 5 we discuss the implications of our work

for digital libraries, and in Section 6 we draw conclusions and outline future work.

2. RELATED WORK

Our study has three areas of related work: literature on interacting with library shelves (Section 2.1), book selection (Section 2.2), and document triage (Section 2.3).

2.1 Interacting with library shelves

Surprisingly little literature exists on how readers interact with library shelves, and the majority of this work focuses on children. One such study is Moore's work that examines children's information search processes by observing children interacting with library shelves [28]. Moore noted that the children often selected the first book they found with a Dewey number related to their query rather than comparing all the books on their topic for usefulness. She also noted that the majority of children could not effectively use the library shelves.

Reutzel and Gali's ethnographic observation of children investigated how children select books from the library shelves [32]. For their study, they developed a *classification of observable book selection behavior*, which we adopt for our study. The observed children were strongly biased toward books on shelves at eye level (60% of selected books), and tended to select books that were in some way eye-catching to them. They were also very aware of the geographical layout of the library and could successfully locate the section, shelf and position of books they were looking for.

Borgman, *et al.* [7] extensively studied the use of the shelf metaphor in a children's DL. The observed children were largely effective at using a DL interface based on the shelf metaphor. This efficacy is in keeping with observations [28, 32] that children seemed to prefer approaching the shelves directly or asking a librarian as opposed to using the catalogue in physical libraries.

Studies of adults' use of library shelves are limited: McKay and Conyers observed that users struggle with every aspect of locating library books on shelves, but provide little insight into how readers select books at the shelves [26]. Stelmaszewska and Blandford's ethnographic study of computer scientists using the physical library noted the occurrence of serendipitous discoveries at the library shelves [35]. This work also showed how participants make judgments about the utility of books based on visual cues such as their apparent age and whether or not they were dusty (a reflection of recent use). In contrast, [8] found that purchasers in bookshops first turned to shop assistants for help, then for the most part engaged in a 'grab and go' strategy, with little browsing or serendipitous discovery.

In summary, the literature on how adults interact with bookshelves to select books is decidedly scant; to our knowledge our study is the first examining this behavior specifically.

2.2 Book selection

There is limited work on how people select books, though some of this work is quite detailed and thus provides significant insight.

Flipping behavior (particularly flipping linearly through books) was reported in a study of readers' engagement with ebooks [25]. [28] and [32] both reported children flipping through books to make decisions, however, similarities end there. Reutzel and Gali observed children selecting and rejecting books based on the presence or absence of images within the book, value statements about the content of the book, and external limitations such as the number of books they were allowed to have out of the library at a time. Moore observed children using textual features such as the table of contents, index and headings. These differences may result from the different nature of the studies; in [32] children selected books freely, in [28] children selected books to meet a pre-defined information need. The table of contents (TOC) was observed to be used by college students when comparing similar books with the intention of meeting a specific information need [36]. This dependence on TOC—and, to a lesser extent, index—was also noted in a number of studies on how searchers find information within books [6, 19, 21, 35], an activity likely to influence book selection. Ooi [30] found evidence that recommendations from popular media and, to a lesser extent, peer recommendations influenced book selections. Similarly, [8] observed that many customers in bookshops used references to non-bibliographic cues to describe the book they needed: for example, popular cultural relationships (such television tie-ins), cover information, reviews in mainstream media, and book appearance.

In summary, the literature shows a range of information being used in book selection. Behaviors used to leverage this data include flipping, skim-reading, reading the TOC, seeking recommendations, and examining the outer book covers. What is little studied is how readers integrate these behaviors with their interactions with library shelves. While [35] reports work on this with computer scientists, no broader study has been conducted.

2.3 Document triage

While there is little information regarding how readers choose between available books, choosing between journal articles has been well studied. This process is referred to in the literature as 'document triage'. [24] and [35] studied document triage in physical journals where readers use the TOC, article titles and authors as well as images and figures (assessed while flipping though the document) to select reading material. Online articles are evaluated in similar ways, with readers using title, author, abstract, introduction and conclusion to evaluate works of interest, and to a lesser extent images and emphasized text [3, 9, 20]. Readers accessed this information by scrolling rapidly through documents [20] (sometimes through and back), the digital equivalent of flipping (though much more prone to leaving users feeling lost [19]). This strong correlation between on- and offline behavior suggests that selection strategies translate well between physical and digital formats, at least for journal articles.

3. METHODOLOGY

Here we describe our methodology for understanding the book selection process in academic libraries. Section 3.1 outlines our observation and interview technique, Section 3.2 describes the target libraries, and Section 3.3 gives participant demographics.

3.1 Observation and interviews

All studies took place in academic or government libraries. We visited libraries individually or in pairs and observed readers' book selection behavior, paying careful attention to the location of sampled books and book sampling behavior. We typically observed interaction with two to five books before briefly interviewing readers about their book selection and motivations in book choice.

We coded observational data using the framework introduced in [32], however we cut the original six criteria down to three objectively observable phenomena: how readers located a desired book within the library, which parts of a book users examined to

determine its usefulness, and where in the library participants looked.

Interviews were executed in situ, and explored how readers made the decision whether to use a book, how they searched for the book, where they would use the book, what they were looking for, what notes they had taken, whether they had considered an ebook, and any other interesting behavior. We also collected demographic information. Due to the low frequency of in-person use for the government library we studied, we typically did not conduct observations in this library, though we interviewed a number of researchers.

The ability to generalize from these results is of course limited by the location of the study: our findings are unlikely to be applicable in a fiction or public general library context.

3.1 Target libraries

All three libraries in this study were academic or government research libraries. At the University of Waikato, New Zealand (UoW), conducted our study in the main library. At the Freie Universitaet Berlin (FUB) Germany, we conducted our study mostly in the Philological library, but also in the Library for Psychology and Education. Our final study location was Bundesamt fuer Materialforschung and -pruefung Berlin (BAM). BAM is the Federal Institute for Materials Research and Testing, and its library functions are those of a specialized research library, and as such we treated it slightly differently (see Section 3.1). Each of these libraries holds specialist non-fiction books (such as reference books, technical books and textbooks). By selecting these libraries, we restricted our study to information access in educational and scholarly contexts (not recreational).

3.3 Participant demographics

We interviewed 65 readers (20 at UoW, 25 at FUB and 20 at BAM), 51 of whom were also directly observed interacting with books in the stacks. Because the labyrinthine layout at BAM made it difficult to track patrons though the stacks, some participants at BAM were interviewed but not directly observed in their interactions. Of the interviewed readers 33 were male and 32 female (22 and 29, respectively, for observed readers). The age distribution is in keeping with expectations for university and government libraries (see Fig. 1).
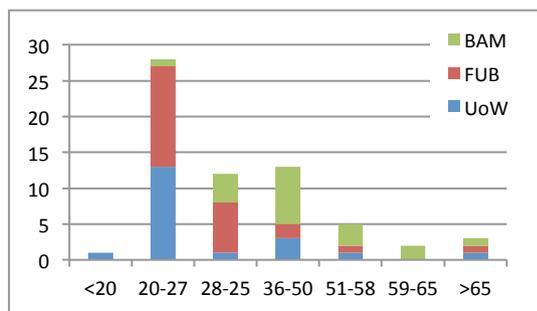
Figure 1. Age distribution of study participants

As would be expected of university libraries, approximately half of the readers interviewed were engaged in research (31) while the other half were conducting their tertiary studies (30). Only two readers were searching for teaching-related material and one patron was visiting the library for recreational purposes, while one declined to respond to that question. Names of participants were neither asked nor recorded; in the remainder of this paper, individual participants are identified by the three letter acronym

for the library in which they were observed or interviewed (BAM, FUB, UoW) appended to a numeric identifier (e.g., UoW07).

Searches covered a wide range of topics, including materials science, anthropology, management, education, and accounting; these topics accord with the research and teaching strengths at the institutions studied. Subjects naturally mirror the specializations of the libraries in which searches were conducted (e.g., German in the philological library at FUB, explosives at BAM).

4. RESULTS

This section summarizes the patterns of behavior observed from our study of participants in the three physical libraries and the results of the brief interviews.

4.1 Physical and digital books

Even in the face of the bias introduced by doing this study, nearly half of all participants (30/65) had considered using an ebook instead of a pint book, and 25 of those expressed a preference for ebooks. Some participants, however, had not considered using an ebook until our questioning suggested it.

Most participants suggested a reason (sometimes quite forcefully) for their preference of one format over another (see Fig.2). 20 participants stated a strong preference for physical books, of whom 13 wanted "real" books, and 7 preferred the "physicality" of paper books (UoW06: "[ebooks] seem so impersonal", UoW04: "print books are easier for research").
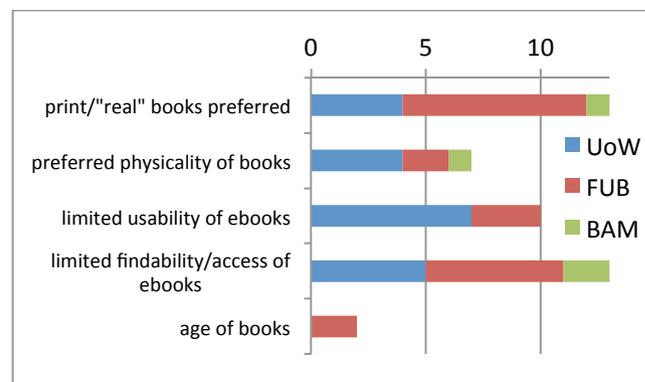
Figure 2. Reasons given for preference for physical books

Ten participants criticized the usability of ebooks in terms of screen access (UoW13: "I get headaches when reading on the screen, Kindle is better"; UoW19: "...gives me a really bad headache"; FUB12: "...strenuous for the eyes") but also in terms of possible interactions (FUB24: "I need to translate [it] and [with a paper copy] I can write the translation into the text"; BAM19: "physical books are better for sharing"). BAM19 felt that physical books enforced "proper reading": "it is important to read everything, not just what comes marked in yellow when you search…the students have to learn that again, to really look at things".

The age of a book was frequently mentioned with respect to both physical and digital books: some participants claimed they could not or would not be able to find an ebook to meet their needs because "only new books are ebooks" (FUB24). Conversely, some participants complained about the age of physical books available in the library FUB02: "I'm annoyed by how old all the books are".

While most readers were aware of the availability of electronic resources (only four participants were not aware of the availability of ebooks), the electronic nature of the materials did not seem to affect the understanding and interaction of some participants. One participant asked, "can you access [ebooks] at home?" (UoW11), and a number of BAM users liked the idea of "books in PDF so you can print them for reading" (BAM13). In both these examples the electronic nature of ebooks is poorly understood at best, and ignored at worst.

It is clear from our questioning about ebooks that physical books are both clearly understood by readers, and hold a real draw for readers. Some of the reasons for this preference became evident on observing readers' interactions with the library as a physical space. We observed several library users sleeping or sprawling, and we saw repeated evidence of the influence of the physical nature of the library on information seeking.

Several readers explicitly named serendipity as a reason for coming to a physical library (FUB14: "I prefer to go to the library and look for books on the shelf to see what you can find serendipitously").

Figure 3: Reading and sorting books on the floor

Some people took chairs into the shelves to read there or to look at the books (3 of 51). A number of readers were found crouching, kneeling or sitting on the floor (13 of the 51 observed readers, see Fig. 3). In part readers did this to reach the books on the lowest shelves, but also to sort selected books on the floor. Readers typically had several books in their hands, not all of which they planned to take out. People sitting on the floor often made piles of books (definitely borrow, maybe borrow) and also moved books around on the floor while making decisions. This behavior mirrors the observations in [32] of children selecting four or five books which they would line up for deeper consideration, deciding which to check out among this smaller subset.

The physical layout was not without its problems, however: if readers did not write down adequate information they would have to leave the stack to repeat their searches (BAM01:"I will have to go back to the catalogue and look again"). Sometimes they forgot to check whether the book was available in the library or out on loan. Some users were confused by local cataloguing systems, rather than finding they supported serendipity ("they [the books] are all in the wrong order - I am completely lost" (UoW10); "I do not understand the system they use here" (FUB 05)). These experiences represent the negative sides of the physical library reported in previous literature (for example [26, 35]).

4.2 Searching the catalogue

34 of the 51 interviewed participants had searched the catalogue before entering the stack; 16 had searched the online catalogue using a library computer (see Fig. 4). Nine participants did not search the catalogue, saying they were familiar enough with the layout of the library that they could go straight to the shelves or sections where books they wished to use were found.

Participants at BAM used the library almost exclusively remotely, searching the catalogue from their offices and then ordering any useful books to be sent over. Many BAM participants claimed, in fact, not to use the library because "everything is on the intranet" (BAM 15)—unaware that this electronic access via the intranet or online is in fact a service provided by the library.

A number of people reported difficulty with the catalogue and the library layout. It seemed that often they expected the catalogue to work similarly to Google (UoW09: "Why does it not work like Google?"), and were disappointed to find it much harder to use.
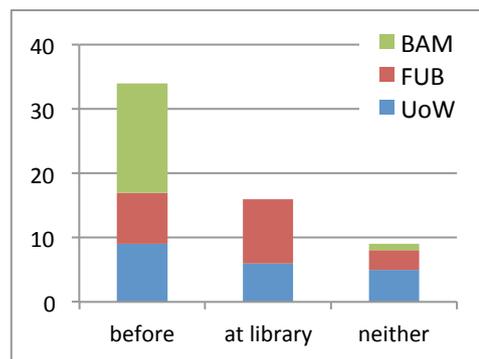
Figure 4: Catalogue use

27 participants reported keyword searching in the catalogue, 22 title searching, 13 authors, and 8 BAM participants used specialist standard identification numbers. 4 participants used other details, including publication date and recommendations from colleagues.
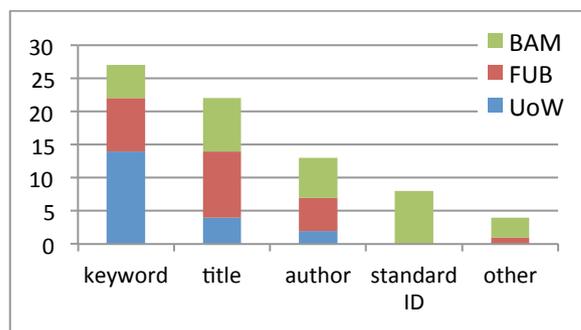
Figure 5. Catalogue search strategies

There were differences between institutions as to the search strategy used (see Fig. 5): UoW participants searched mostly by keyword (only 6 of 20 used title or author), whereas FUB participants mostly (20 of 24) used title or author. These differences may be due to local variations in catalogue software, or they may reflect the different participant groups seen at FUB and UoW: UoW participants were younger and more likely to be studying (as opposed to conducting research) than those at FUB. BAM participants, like FUB participants typically did not search by byword (only 5 of 22 did). BAM researchers are typically experienced in their field, and reported using the library only when they had identified a resource that they wished to access.

4.3 Artifacts used as finding aids

Participants who searched the catalogue before entering the library stacks often had notes to assist in locating books. 24 participants permitted the researchers to photograph their notes (3 refused). Few BAM notes were recorded as the participants had largely discarded the notes before being interviewed, demonstrating that these notes are used to store information only temporarily while readers navigate to desired information resources on the shelves.
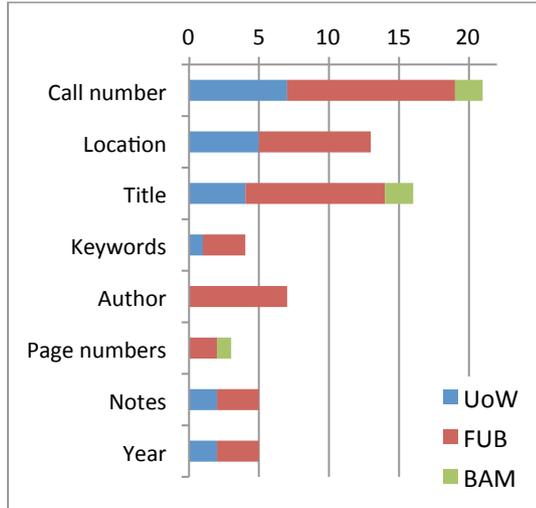
Figure 6. Details included on notes created to assist in locating books
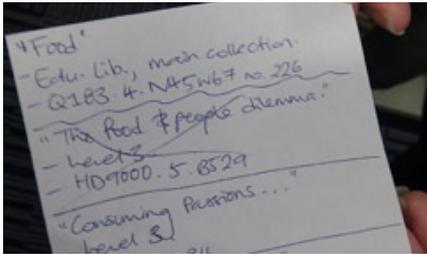
Figure 7. Note to assist in locating books (UoW19)
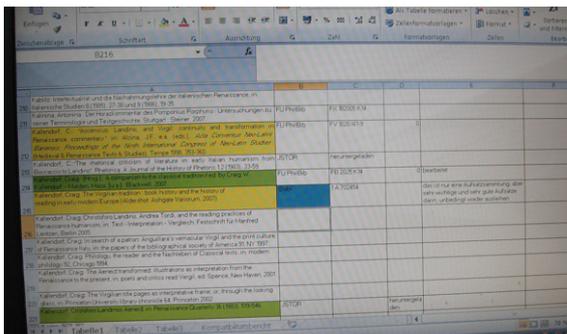
Figure 8. Electronic notes (FUB19)

Of the 24 photographed notes, 21 were handwritten, two were printed and one participant had brought a laptop to the library. Notes included call numbers, location, and titles (see Fig. 6). Figure 7 (UoW19) shows a typical example of a note. The participant had crossed out books they had checked but rejected as not suitable.

One participant (FUB19) used an electronic spreadsheet to keep track of their research material. The list included both printed books and electronic material (see Fig. 8); color-coding was used to indicate importance and to keep track of what had been read or examined.

4.4 Looking at the shelves

Readers located potentially interesting books on the shelves in one of four ways (see Fig. 9). The first strategy was "scanning a section" by looking along several shelves belonging to the same topic, the second was "sampling the shelf" in a more concentrated view (sometimes moving in an orderly top-to-bottom and left-to-right pattern, sometimes taking a more scattershot approach such as for display books), and the third was standing in one place close to the shelf and examining the shelf above, at or below eye level. Finally, some readers found books on display shelves of newly acquired or other promoted books (called "display books" hereafter). The relatively infrequent use of display books is commensurate with the small proportion of any library's collection on display at any time.

While three readers discovered books on display shelves, they were not the only readers to discover books in passing. One reader walked past a shelf and noticed a book's cover, which could be seen at the end of the shelf (it turned out to not be relevant to the topic of interest). Three of the four people who "found" books in passing were on their way to/from the washroom. In addition to these four, several participants found books by scanning the shelves adjacent to those containing a book of interest. This finding supports participants' assertions that the physical library facilitates serendipitous discovery of books.
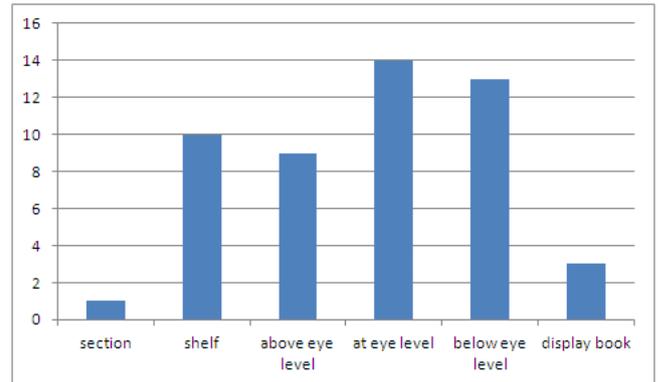
Figure 9. Sampled book geography

4.5 Choosing books from the shelves

51 readers were observed choosing books from the shelves. Their actions in doing so are described in Figure 10. Readers touched books (without pulling them out), tilted them (typically to look at the cover), pulled or half pulled books (and put them back). Only 31 out of the 51 took books completely off the shelves for further examination. When examining books on the shelf, readers cocked their heads to read what was on the spine, and two mumbled as they were looking for specific books. The use of the spine refers to readers who were observed to explicitly read the spine; only 10 of 51 readers did this. As described in Section 4.4, four readers chanced upon books displayed in such a way that their covers were already visible. All other readers had to base their decision to investigate further on the spine of the book.

Tilt, pull and half pull were interpreted as an intention to glimpse the cover when this could not be examined directly. "Pull" refers

to the action of lifting the book off the shelf to examine the front cover without opening the book, whereas 'take off shelf' refers to removing a book and then opening it to investigate further.
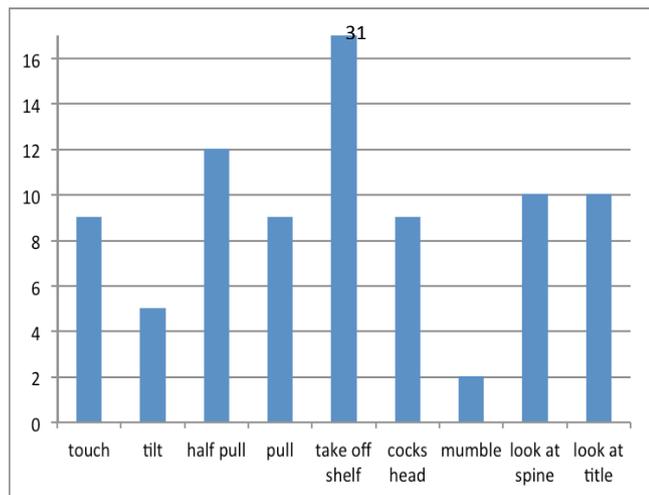
Figure 10. Book search behavior

4.6 Evaluating books

31 readers examined books more closely than just looking at the shelves in their decision making; some examined books in the stacks, while some removed their selections to a nearby table to make a decision. We observed more than half of all readers looking at the TOC, many of these followed up by reading sections of pertinent chapters and/or flipping through the book (see Fig. 11, blue). A number of readers were observed flipping *backwards* through books. Overall four main strategies were used for sampling the content of books: examination of the front cover, flipping backwards or forwards through the book reading part of the book, and examining the TOC.
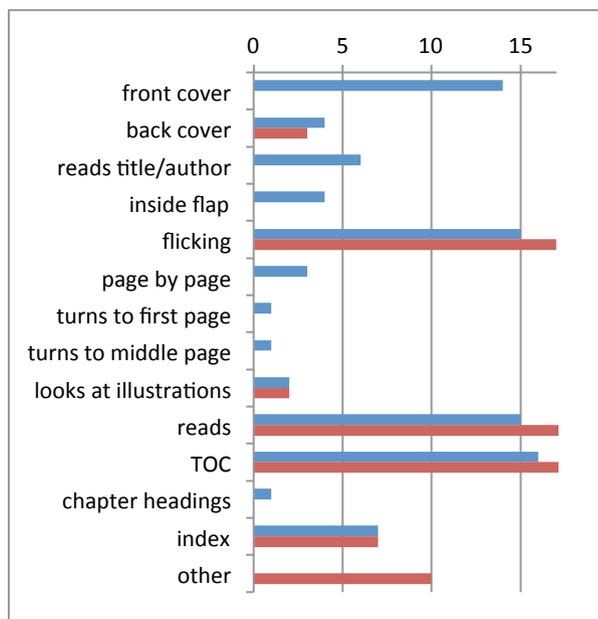
Figure 11. In-book browsing behavior (blue) and reported decision-making behavior (red)

Only one visitor carried out a "grab and go" behavior— he had returned for the very same book that he had used that morning. This is in stark contrast to Buchanan's bookshop observation where over half of the observed bookshop readers showed a "grab and go" strategy [8]. Buchanan's study, however, reports behaviors of those who had already asked a question of a bookseller and who were often purchasing for others, where our participants did not report talking to a librarian and were looking for books for their own use.

Readers' self-reported decision making (triage) behavior was largely in accordance with observed book sampling behavior (see Figure 11, red). Participants reported placing importance on the TOC and reading pertinent parts of a book as influencing their decision making behavior. Many of them also believed that flicking through the book was important. Participants who looked at the index and at illustrations reported using them for decision making. Participants used the back of the book to determine content and target audience. 'Other' facets mentioned were references and edition. Two readers needed to check whether the book was mentioned by their lecturer. One person said they wanted to spend time with a book (about 10min) to see if it "felt like" the right text.

4.7 Using books

Participants planned to access books in three ways (see Fig. 12): using the books in the library (15), checking the books out of the library to use at home or in the office (25), or copying all or part of the book to use later (20). Readers' choices were naturally strongly influenced by the options on offer at each library. All three libraries offered reading desks; however, the ones at BAM were more geared towards brief interactions with the books close to the shelves, whereas FUB and UoW offered work areas (partially equipped with computers and LAN connections for laptops). BAM and UoW had copy machines that allowed hardcopy as well as electronic copy into PDF documents. Copy machines at FUB only allowed hardcopy.
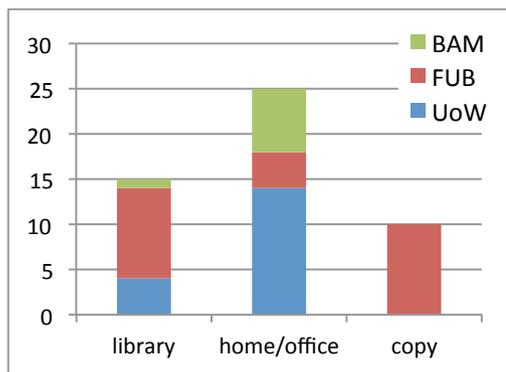
Figure 12. Planned interaction with selected books

Researchers at BAM reported that they visit the library to confirm the validity of citations: upon finding the relevant books they immediately turned to the pages given in the citations, to check the text they expected was present in the books, and if it was indeed present then they made copies to read later in their offices. The FUB philological library is predominantly a reference library (that is, most books cannot be taken out of the library) and so it provides desks and computing facilities in the library. Consequently, readers either used books in the library or copied out the parts they were interested in. In contrast, UoW offers a relatively small number of workspaces, and so UoW patrons check books out or copy them and take copies home (though

explicit warnings about copyright infringement over photocopy machines and the cost of copying may influence preferences in this behavior).

5. Discussion

Drawing on summaries of observations and interviews (Section 4), and in reference to previous studies, we here explore interface and interaction design implications for ebook collections.

5.1 Finding books on the shelves.

Physical bookshelves are awkward to browse: only the tallest patrons are able to easily scan the top-most shelves, and for all but the shortest, viewing the lowest shelves involves bending or kneeling on the floor (Section 4.1). In our study these physical limitations are reflected in the participants' shelf scanning patterns: shelves above eye level were scanned less frequently (Section 4.3) than shelves at eye level or lower. This pattern of behavior is even more pronounced in children [32].

We might expect the digital display of ebooks to ease the tyranny of height imposed by shelves in the physical library; however, the conventional list-based display of digital library documents introduces similar difficulties. In digital library, ebooks are no longer less accessible due to the physical limitations of readers and shelves; instead, the digital library interface typically limits the collection view to a one-dimensional list, with only a few (often 10 to 25) ebooks visible at a time. The facility to scan hundreds of books on a physical shelf in seconds has been reduced to slowly limping from one keyhole view to another in a digital library – suggesting the need for further research into more quickly and more easily browsable DL displays. For example, Shneiderman suggested presenting digital library results as color-coded dots in a 2D-space [34]. Another option is to show ebooks in a two-dimensional space representing search similarity (as typically done in catalogues and DLs) and category surroundings (as on a library shelf), one example of which is seen in [7]. The link between searching and browsing could be strengthened by presenting search results with the hits highlighted within their category.

Physical books are typically grouped on shelves by subject classification (e.g., Dewey, Library of Congress). Readers arrive at relevant books either through searching the online catalogue (See Section 4.2), or browsing the shelves of a relevant subject (where that subject and its physical location in the library may have been identified in a previous online catalogue search). Browsing and searching are thus interleaved, though somewhat awkwardly, in the physical collection and online catalogue. While we might expect that this interleaving would be easier in a purely digital library, this may not be the case. In the popular Greenstone digital library system, for example, it is not possible to search within a subject classification, and not easy to switch to browsing the subject classification of a document identified by a search [27]. The DL users are thus less likely then physical library users to arrive at an ebook within a subject display, and so an ebook is most frequently viewed as part of the search term list (book among other books matching that search term) rather than the physical library view of a book in context with other books in the same subject category. At the same time, participants in our study reported valuing browsing by subject context (Section 4.1) and generally seemed to have a sound understanding of how the shelf classification worked.

The presentation of books in a physical library remains largely constant. The extent to which readers rely on this stability is documented by the observed nine library visitors who approached the shelves directly because they "know where things are" (FUB03). The layout allows visitors to build a mental map of "their" part of the library. This is not possible when ebooks are viewed primarily as elements within search results; search result displays are transient and it is notoriously difficult to recall previous search terms. Stronger support for subject-related access may help readers to grow familiar with an ebook collection.

Physical bookshelves let humans do what we do well: quickly scan and identify what might interest us from a herd of similar objects [5]. Our study confirms that people scan shelves, or sections or parts thereof; rarely did a reader look only at the book they wanted. Serendipitous finds of books were observed (Section 4.4) and valued by participants (Section 4.1); they typically happen when readers physically encounter books as they move though the library [26, 35]. In ebook collections, limited serendipity can be achieved by adding more stable displays (e.g., by using the 2D category references discussed earlier), however, there is further scope for offering additional ebook browsing to online users. Currently there are two documented ways to offer serendipity in digital libraries: the highlighting of a single, featured document, and streaming collage display [4]. The use of a single featured document does not meet the need for serendipity described by participants in this and other studies, as it shows only a single document as opposed to groups of related documents. Streaming collage displays have been demonstrated to be highly effective for collection understanding, which suggests they may also be effective for serendipitous discovery [12].

5.2 Selecting books from the shelves

Adult readers were observed to predominantly judge books by sight (Section 4.5), not needing to run fingers along the spines as was common for children [32]. The tilt and look (pull and half pull) mechanism was prevalent in our results as was the sampling act of opening, flicking through and reading (Section 4.6), whereas children tend not to read or skim within books. The prevalence of physical interaction with books during the selection process would suggest that it is worthwhile investigating interfaces that enable people to quickly, easily, and intuitively 'touch' a book. 'Touching' ebooks is increasingly possible with tablet technologies, and could result in the display of metadata, a selection of images, or other document navigation devices. It is clear from our study that touch is a common and useful way to interact with objects, and being able to do that in a digital space could bring people closer to "get a feel" for the documents (Section 4.4).

All observed participants at some point interacted with books to discover more information than was visible from the spine. The large number of readers considering the covers of books in the physical library suggests that information and metadata contained on this artifact is germane to the book selection process and thus also of significance to the presentation of books in a digital library. Metadata found on the covers of books within academic libraries commonly includes titles, authors, book synopses or book reviews. Thus a quick pull or half-pull can reveal information about the tone and target audience of a book. DL and catalogue interfaces do not offer the same visual cues (often no cover image is provided); presenting this type of visual data in DLs is likely to benefit information seekers.

Several studies [8, 26, 35] observed the use of non-conventional cues such as book age, size and dustiness, cover image, and popular cultural associations in book decision-making. Our study

corroborates these findings, as participants used age, cover, and images for decision making (Section 4.2). These types of cues are typically absent from ebook collections, and as such represent a significant opportunity to develop more usable DLs.

5. 3 Evaluating and comparing books

Participants quickly scanned significant portions of books by flicking through the pages (Section 4.6). Flicking can give an overview of the contents of a book and an indication of a book's style. Flicking and quick reading as decision making behavior may not translate directly into current ebook environments. The act of opening an ebook often constitutes 'lending' of the book. This distinction may not matter for libraries that buy bundles of ebooks, but for libraries that have a pay-per-view system or use-strategy for buying an ebook once a number of readers have downloaded the book, flicking behavior could have serious budgetary consequences. An alternative approach to linear digital flicking could be the provision of a field of thumbnails for all the pages of the book. Additionally one interviewed reader commented that he "rather flips than scrolls" (BAM09)–even though his comment was meant to justify a preference for physical books over ebooks. One possible implementation may be a swiping motion across pages as in [19] instead of scrolling though ebook—shown in to be more effective for many users in the aforementioned study.

Readers in the physical library were observed stacking, ordering and closely inspecting books at various stages in their decision making process, often on the floor, or by re-purposing an empty shelf or ledge (Section 4.1). There is no analogue for this in DLs, and as mentioned above budgetary constraints may even act as a constraint on this behavior. Offering readers a lightweight way to compare ebooks side-by-side in DLs would likely significantly improve users experience significantly, and as such is a ripe avenue for future work.

5.4 Keeping track of books

Participants planned to read at the library, copy or take books out of the library (Section 4.5). The problem of how to keep track of what has been evaluated, read or identified to be read is encountered for physical and digital books alike. A physical book checked out of a library is a physical reminder that the book is relevant. However, once the book is returned, the reminder is lost, and any records about books that have been read or remain to be read must be kept by readers themselves. Some participants used written notes to cross out books that they had evaluated and dismissed. These notes were typically discarded in the library; only one patron was observed to keep track of books they wanted to read in future (keeping an electronic reference, Section 4.7). This lack of a personal book interaction history can waste readers' time: several participants reported mistakenly re-evaluating books they had previously examined and discarded as irrelevant.

Often readers found several books of interest during the information seeking process that they want to set aside for later reading (a discovery also seen in [35]); how readers manage this bears further investigation. This is an area where DLs and electronic books have an advantage. Though currently not a main feature of available e-readers and DLs, it would be a simple extension to highlight books a reader has accessed, read or plans to read—or has rejected as not relevant.

The physical library helps people keep track of books during the selection process by providing context: readers often elected to triage books within the shelves rather than moving to a more comfortable location (Section 4.1). This context allows readers to readily scan possibly useful books, and cycle rapidly between examining individual books and scanning shelves. This facility for context is an advantage of physical libraries not seen in DLs; the ability to provide stronger contextual information is an avenue for future DL design work.

In physical libraries we have several cues when books are added to the collection: for example, frequently acquisitions will first appear in a 'new books' display. A reader familiar with a particular section of a library can tell at a glance when new and unfamiliar volumes are added to 'their' shelf. Identification of new results to prior searches can be implemented in ebook collections with relative ease by alerting services, resulting in personalized recommendations to readers [16].

6. Conclusions and future work

The study presented here is a pioneering look at the book selection process as it occurs at the library shelves. Our findings are based on a combination of observation and interviews with library users. Our discoveries have a variety of implications for DLs, particularly in terms of lightweight interaction, browsing, and decision-making metadata. This study also provides considerable scope for future work.

Even though it appeared that sorting books and document triage would be more comfortable at a near-by desk, participants largely preferred to work proximate to the shelves even if that meant they had to sit on the floor (Section 4.10). Here proximity retains context in the process of looking through a book shelf or section. Ebook collections can allow the capturing of richer contexts than is possible in the physical library. Context in a physical library can only reflect current interaction history, whereas DLs can potentially maintain records of previous interactions with the collection.

It is evident that many of the physical features of library space support users in their book decision making. The escalation from a glance to reading the spine of a book to examining the cover to flipping through the pages is simple and intuitive to readers, as is the ability to compare books side by side. Conversely, the physicality of library spaces also cause difficulties for users: most shelves are too high or too low for most users and none of the libraries in question provided adequate sorting space, meaning users often sorted books on the floor. There is a real opportunity for DLs to replicate the positive aspects of physicality by allowing users to 'touch' books in lightweight ways and to compare books side by side without imposing the burdens of physical space.

Related to the issue of physical space, though not dependent on it, is the way in which physical libraries facilitate browsing. Many users in our study deliberately used the co-location of related books to assist their information seeking, and for others it facilitated serendipitous discovery of useful resources. Browsing as an activity is poorly supported in DLs, but 'co-locating' works which are in some way related to facilitate serendipity is possible. Again, DLs have an advantage over physical environments in that it is possible in a DL to 'rearrange the shelves' according to an information seeker's interests.

Finally, we noted that information seekers used a wide range of criteria to make decisions about books. Some of these were the type of bibliographic metadata typically available in DLs (such as author and title), but some were less widely available, such as cover image, usage history, images within the book, and table of contents. To best facilitate decision making, DLs must seamlessly provide this information to readers at the point of need.

This work also raises a number of questions for future examination. Our work shows considerable differences from earlier work with respect to the number of books examined from above and below eye-level; eye tracking studies would help us to understand what drives these behaviors. Our work also raises the question of which decision-making criteria are already used when choosing ebooks. A careful analysis of log data would explore this issue.

Our work, while raising many questions, is an examination of a largely unstudied part of the information seeking process: the act of selecting among books at the shelves. This work has considerable implications not just in the physical realm, but also for digital libraries. Based on our work, the three main challenges for DLs in this area are supporting serendipitous discovery and browsing, providing a lightweight means of investigation and comparison of books, and supplying the information users actually want to use to choose between books. Meeting these challenges would be a significant step to replicating the advantages of physical libraries in the online sphere and providing the usable, useful DLs that readers so clearly desire.

9.References

[1] Adler, A., Gujar, A., Harrison, B. L., O'Hara, K. and Sellen, A. A diary study of work-related reading: design implications for digital reading devices. In *Proceedings of CHI '98* Los Angeles, California, United States. ACM Press/Addison-Wesley Publishing Co., New York, NY USA,1998. 241-248.

[2] Bae, S., Hsieh, H., Kim, D., Marshall, C. C., Meintanis, K., Moore, J. M., Zacchi, A. and Shipman, F. M. Supporting document triage via annotation-based visualizations. *ASIST Proceedings*, 45, 1, 1-16.

[3] Bae, S., Marshall, C. C., Meintanis, K., Zacchi, A., Hsieh, H., Moore, J. M. and Shipman, F. M. Patterns of reading and organizing information in document triage. *ASIST Proceedings*, 43, 1, 1-27.

[4] Bainbridge, D., Cunningham, S. J. and Downie, J. S. Visual Collaging of Music in a Digital Library. In *Proceedings of ISMIR* Barcelona, Spain, October 10-14, ,2004. 397-402.

[5] Bates, M. J. The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, 13, 5, 407-424.

[6] Berg, S. A., Hoffmann, K. and Dawson, D. Not on the Same Page: Undergraduates' Information Retrieval in Electronic and Print Books. *The Journal of Academic Librarianship*, 36, 6, 518-525.

[7] Borgman, C. L., Hirsh, S. G., Walter, V. A. and Gallagher, A. L. Children's searching behavior on browsing and keyword online catalogs: The Science Library Catalog project. *JASIS*, 46, 9, 663-684.

[8] Buchanan, G. and McKay, D. In the Bookshop: Examining Popular Search Strategies. In *Proceedings of JCDL 2011* Ottawa, Canada, June 14-17. ACM, New York, NY, USA, 2011.

[9] Buchanan, G. and Owen, T. Improving skim reading for document triage. In *Proceedings of IIiX '08* London, United Kingdom. ACM, New York, NY USA, 2008. 83-88.

[10] Cousins, S. A. In their own words: an examination of catalogue users' subject queries. *J. Inf. Sci.*, 18, 5, 329-341.

[11] Crabtree, A., Twidale, M. B., O'Brien, J. and Nichols, D. M. Talking in the Library: Implications for the Design of Digital Libraries. In *Proceedings of DL 97* Philadelphia, PA, USA. ACM Press, New York, NY USA, 1997. 221-228.

[12] Cunningham, S. and Bennett, E. Understanding Collection Understanding with Collage. In *Proceedings of ICADL* Bali, Indonesia, December 2-5. Springer Berlin / Heidelberg, 2008. 367-370.

[13] Cunningham, S. J., Knowles, C. and Reeves, N. An ethnographic study of technical support workers: why we didn't build a tech support digital library. In *Proceedings of JCDL 01* Roanoke, Virginia, United States, June 24-28. ACM, New York, NY USA,2001.

[14] Cunningham, S. J., Reeves, N. and Britland, M. An ethnographic study of music information seeking: implications for the design of a music digital library. In *Proceedings of JCDL 03* Houston, Texas. IEEE Computer Society, Washington, DC, USA ,2003. 5-16.

[15] Greifeneder, E. The Impact of Distraction in Natural Environments on User Experience Research In *Proceedings of TPDL '11* Berlin, Germany, 26-28 September. Springer, Berlin/Heidelberg ,2011.

[16] Hinze, A., Schweer, A. and Buchanan, G. An Integrated Alerting Service for Open Digital Libraries: Design and Implementation In *Proceedings CoopIS 05* Agia Napa, Cypris. Springer Berlin / Heidelberg, 2005.

[17] Kuhlthau, C. C. Inside the Search Process: Information Seeking from the User's Perspective. *JASIST*, 42, 5, 361-371.

[18] Lau, E. P. and Goh, D. H.-L. In search of query patterns: A case study of a university OPAC. *Information Processing & Management*, 42, 5, 1316-1329.

[19] Liesaputra, V. and Witten, I. H. Seeking information in realistic books: a user study. In *Proceedings of JCDL '08* Pittsburgh PA, PA, USA. ACM, New York, NY USA,2008. 29-38.

[20] Loizides, F. Investigating Document Triage on Paper and Electronic Media. In *Proceedings of ECDL 2007*. Springer Berlin / Heidelberg,2007. 416-427.

[21] Malama, C., Landoni, M. and Wilson, R. Fiction Electronic Books: A Usability Study. In *Proceedings of ECDL '04*. Springer Berlin / Heidelberg, ,2004. 69-79.

[22] Marchionini, G. *Information Seeking in Electronic Environments*. Cambridge University Press, Cambridge, UK, 1995.

[23] Marshall, C. C. *Reading and Writing the Electronic Book*. Morgan & Claypool, Chapel Hill, NC USA, 2010.

[24] Marshall, C. C. and Bly, S. Turning the page on navigation. In *Proceedings of JCDL'05* Denver, CO, USA. ACM, New York, NY USA, 2005. 225-234.

[25] McKay, D. A jump to the left (and then a step to the right): reading practices within academic ebooks. In *Proceedings of OZCHI 2011* Canberra, Australia. ACM, New York, NY USA ,2011. 202-210.

[26] McKay, D. and Conyers, B. Where the streets have no name: how library users get lost in the stacks. In *Proceedings of CHINZ 2010* Auckland, New Zealand, 8 July. ACM, New York, NY USA ,2010. in press.

[27] McKay, D., Shukla, P., Hunt, R. and Cunningham, S. J. Enhanced browsing in digital libraries: three new approaches to browsing in Greenstone. *JoDL*, 4, 4, 283-297.

[28] Moore, P. Information Problem Solving: A Wider View of Library Skills. *Contemporary Educational Psychology*, 20, 1, 1-31.

[29] Nordlie, R. "User Revealment"— A Comparison of Initial Queries and Ensuing Question Development in Online Searching and in Human Reference Interactions. In *Proceedings of IR 99* Berkeley, CA, USA, August 15-19. ACM Press, New York, NY USA ,1999. 11-18.

[30] Ooi, K. How Adult Fiction Readers Select Fiction Books in Public Libraries: A Study of Information Seeking in Context. Masters, Victoria University of Wellington, Wellington, New Zealand, 2008.

[31] Pearson, J., Buchanan, G. and Thimbleby, H. HCI design principles for ereaders. In *Proceedings of Booksonline '10* Toronto, ON, Canada. ACM, New York, NY USA, 2010.

[32] Reutzel, D. R. and Gali, K. The Art of Children's Book Selection: A Labyrinth Unexplored. *Reading Psychology*, 19, 1, 3-50.

[33] Rowlands, I., Nicholas, D., Jamali, H. R. and Huntington, P. What do faculty and students really think about e-books? *Aslib Proceedings*, 59, 6, 489-511.

[34] Shneiderman, B., Feldman, D., Rose, A. and Ferre Grau, X. Visualizing digital library search results with categorical and hierarchical axes. In *Proceedings of DL '00* San Antonio, Texas, United States. ACM, New York, NY USA,2000. 57-66.

[35] Stelmaszewska, H. and Blandford, A. From physical to digital: a case study of computer scientists' behavior in physical libraries. *JoDL*, 4, 2, 82-92.

[36] Stieve, T. and Schoen, D. Undergraduate Students' Book Selection: A Study of Factors in the Decision-Making Process. *The Journal of Academic Librarianship*, 32, 6, 599-608.

[37] Wallace, P. M. How do patrons search the online catalog when no one's looking? Transaction log analysis and implications for bibliographic instruction and system design. *RQ*, 33, 2, 239-253.