



McKay, D. (2011). A jump to the left (and then a step to the right): reading practices within academic ebooks.

Originally published in *Proceedings of 'Design, culture and interaction', the 23rd Annual Conference of the Australian Computer-Human Interaction Special Interest Group (OZCHI 2011), Canberra, Australian Capital Territory, Australia, 28 November-02 December 2011* (pp. 202-210). New York: ACM.

Available from: <http://dx.doi.org/10.1145/2071536.2071569>

Copyright © ACM, 2011. The definitive version was published in *Proceedings of OZCHI, (2011)*.

This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. If your library has a subscription to these conference proceedings, you may also be able to access the published version via the library catalogue.



A jump to the left (and then a step to the right): Reading practices within academic ebooks

Dana McKay

Swinburne Library & The Swinburne Institute
Swinburne University of Technology, PO Box 218 John Street, Hawthorn
VIC 3122, Australia
dmckay@swin.edu.au

ABSTRACT

Considerable attention has been paid to how readers find, triage, navigate and read periodical material such as journal articles. Until recently however, applying these questions to books has been impractical or impossible. This paper reports an exploratory log analysis of ebook usage in an academic library. This study investigates raw usage, document triage practices, and in-book navigation.

Author Keywords

Ebooks, in-document navigation, information use, books, reading, document triage, information behaviour

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI); Miscellaneous.

INTRODUCTION

Reading is undergoing significant change. The use of ebooks has soared in recent months; ebook sales on Amazon surpassed print book sales for the first time in May 2011 (Hamblen 2011). Library ebook usage is also increasing; one study shows a 3-5 fold increase in use even as long ago as the period 2002-2004 (Bailey 2006). In 2007, half of all of all survey respondents at UCL reported having read at least one ebook (Rowlands et al. 2007); in 2008 58% of those surveyed at the University of California reported ebook use (Li et al. 2011).

When users are questioned about their use of ebooks, they cite convenience, portability, searchability and currency as factors in choosing ebooks over print (Heron et al. 2007; Li et al. 2011; McKay 2011). Ebooks are not without their problems, though: users are frustrated by publisher restrictions on use (such as limitations on printing, copying and downloading (Shelburne 2009; Li et al. 2011)). In addition to the difficulties with ebooks themselves, ebook reading technologies are problematic for users. A trial of Kindle ebook readers in an academic environment showed that many users found them unworkable for academic reading (Thayer et al. 2011), and broader studies of ebook readers have shown significant usability problems (Malama et al. 2004;

Pearson et al. 2010; MacFayden 2011). The most recently available study, however, shows that despite their faults, ebook readers are preferable to computer screens for reading ebooks (Li et al. 2011), largely due to their portability. The problems with computer screens are broadly similar to those of ebook readers: using either technology, browsing is difficult (Berg et al. 2010), annotation is challenging at best (Marshall 2010), and the lack of physical cues renders within-document navigation cumbersome (Liesaputra et al. 2008).

Some work has gone into comparing online reading experiences with their paper-based counterparts. Early work focused on the ergonomic aspects of online reading (Dillon 1992); later work investigated reader experience (O'Hara et al. 1997). In addition to this comparative work, some studies have attempted to capture naturalistic print document reading practices in the hope of better informing the design of online reading systems (Adler et al. 1998; Marshall et al. 2005).

The vast majority of work on reading, however, has focused on the reading of scholarly articles and periodicals, perhaps because these are easier to study, and as Marshall points out it is very difficult to study the use of books (Marshall 2010). Only two experiments have assessed how readers engage with print books (Liesaputra et al. 2008; Berg et al. 2010); these studies were both artificial studies of specific information seeking tasks. Recent work in the digital libraries field has shown that users perform differently in a natural test environment than during lab testing (Greifeneder 2011), meaning that even these studies may tell us less about reading than we would hope. Despite the aforementioned studies and the common nature of scholarly book reading (George et al. 2006; Tenopir et al. 2009), we have little idea how readers actually use scholarly books.

This paper presents an exploratory transaction log analysis of usage of a large ebook collection provided by the library at a small, research-active university in Australia. This work was undertaken with the specific intention of capturing naturalistic interaction with books, and is, to the author's knowledge, the first study of its kind.

This paper is divided up into five sections: study background; a description of the methodology used; results of the study; discussion of the results and their implications for design of online reading experiences, and finally future work and conclusions.

OZCHI 2011, Nov 28 – Dec 2, 2011, Canberra, Australia.

Copyright the author(s) and CHISIG

Additional copies are available at the ACM Digital Library (<http://portal.acm.org/dl.cfm>) or ordered from the CHISIG secretary (secretary@chisig.org)

OZCHI 2011 Proceedings ISBN: x-xxxxx-xxx-x

BACKGROUND

There is a range of literature relevant to this work, including investigations of in-document navigation and document triage, work on the use of ebooks, examinations of reading practice, and the (somewhat scant) literature on how readers find books.

Ebooks

The literature on the scholarly use of ebooks is relatively limited, and for the most part based on data two to three years old.

There have been a number of library-based studies to determine how widely ebooks are used, see for example (Littman et al. 2004; Christianson 2005; Christianson et al. 2005). The results of these studies vary, with some claiming greater use of ebooks than print books, and some claiming the opposite. Common among these studies, though, is the finding that discipline affects ebook use, and that ebooks and print books are used under different circumstances.

Various methodologies have been used to gauge reader experience and understanding of ebooks including interviews (Hernon et al. 2007) focus groups (McKay 2011) and surveys (Rowlands et al. 2007; Shelburne 2009). The largest published survey generated 2256 responses from faculty and students at the University of California (Li et al. 2011). These studies have universally found that users like ebooks for their convenience and searchability, however they bemoan poor annotation capabilities and cumbersome navigation. Interestingly, one feature which is mentioned as discomfitting in ebooks is turning the page (Pearson et al. 2010). There is some evidence that page turning disrupts reading even in print (Marshall et al. 2005), suggesting some of the usability problems with digital documents have analogues in more traditional technologies.

The findings from surveys and interviews are supported by empirical studies of information finding in ebooks, which describe poor performance of technology and user dissatisfaction (Liesaputra et al. 2008; Berg et al. 2010).

What it means to read

Pearson makes the point that books are an old technology that has evolved along with reading and they are thus uniquely suited to it (Pearson et al. 2010). This sentiment is echoed by a study of reading in a natural setting, which noted that readers' interactions with paper are second nature, and yet somehow central to the reading experience (Marshall et al. 2005). Despite Marshall's 2005 work, and perhaps because reading is so ingrained in our society reading is hard to study; Marshall pointed out in her 2010 book on ebooks that 'observing people read is by its nature creepy' (Marshall 2010). Nevertheless, there is some work on the nature of reading, particularly in the academic realm.

The nature of reading for work or scholarly purposes was investigated in 1998 by Adler et al.; they identified a range of reading practices. These practices include skimming (even during leisure reading, over 50% of the pages of a periodical may be skimmed (Marshall et al.

2005)), document triage (determining how useful a document is (Marshall et al. 2005)), deep reading for learning and deep reading for editing. In Adler's study and numerous others, for example (O'Hara et al. 1997; Blustein et al. 2011; Thayer et al. 2011), annotation and writing were found to be a part of the majority of work-related and academic reading.

While reading as an activity is unlikely to have changed significantly in recent years, the patterns of what is read in academia have altered dramatically in the past three decades (Tenopir et al. 2009). Scholars are reading more, using the library more, and sourcing and citing documents available online more frequently. Nonetheless, this study and many others (for example (Talja et al. 2003; George et al. 2006)) still show readers using a mix of materials including both scholarly articles and monographs, though the split between these varies by academic discipline.

The author's own recent work in this area has shown that users search for and anticipate using books and articles differently from one another (McKay 2011; McKay et al. 2011). This disparity means that findings on actual use patterns of scholarly articles, web pages or other materials may not be applicable to books.

How readers find books

As Rowlands notes the literature on how users search for and find books is somewhat limited (Rowlands et al. 2007). Key themes in the literature on book finding include reference popular culture (for leisure reading (Buchanan et al. 2011)) and recommendations (Rowlands et al. 2007; Ooi 2008), a theme also seen in the literature about article selection (Talja et al. 2003; George et al. 2006). The work on ebook searching reports that the library catalogue and general web searching are the most common strategies for finding books (Rowlands et al. 2007; Li et al. 2011); conversely, the literature also shows that including ebooks in the catalogue increases use of those ebooks (Rettig et al. 2003; Christianson 2005).

How readers choose documents

How readers select documents of interest, a process also known as 'document triage' (Marshall et al. 2005) has been widely studied for shorter documents (Bishop 1998; Badi et al. 2006; Bae et al. 2006; Loizides 2007). These studies have consistently found that—online or in print—titles, abstracts, section headings, emphasized text and document conclusions are valuable for determining the usefulness of a document. One study found previous knowledge of the topic area is also brought to bear on the decision making process (Marshall et al. 2005), at least for leisure reading, a finding which is likely to be more broadly applicable.

The findings on how users select books are fairly similar to those for periodical materials (Malama et al. 2004; Stieve et al. 2006; Liesaputra et al. 2008; Berg et al. 2010). The major difference is that in books the table of contents is a key piece of decision-making information: undergraduate students will use table of contents and book structure to choose between books (Stieve et al.

2006). Another study found that visual properties of the book—age and how long it appeared to have sat on the shelf—affected decision making (Stelmaszewska et al. 2004), but this implicit information is not available for ebooks and thus cannot affect decision making.

How readers navigate within documents

It is reasonable to assume that readers begin reading fiction books at the beginning and read sequentially from there; the narrative structure of novels essentially requires this. Academic texts such as those in our ebook collection however, are more analogous to journal articles or other texts with clear in-document anchor points including title, abstract, authors and section headings. Nonetheless, given that it is a near-impossibility to study users navigating within any kind of book (Marshall 2010), the literature on in-book navigation is limited. One study compared navigation within books in four formats (three online formats plus print) and found that users were disoriented and could not determine the size of documents online (Liesaputra et al. 2008). Users in this study were also observed marking places in physical books with their fingers, even though post-it notes were readily available, demonstrating the ingrained nature of book navigation. Ingrained navigation of books (and confusion when using ebooks) was also seen in the only other study located on the topic (Berg et al. 2010); in this study users were readily able to search physical books for information, but less so online books.

The patterns of online book navigation seen in these two studies are reflected in studies of online documents: the most commonly mentioned patterns include navigating stepwise through a document, examining the first page only, examining the first and the last page (Alexander et al. 2008; Loizides et al. 2009), and working through a document to the end then working back. Similar to this last pattern is the practice seen in a study of users scrolling back and forth within an online document (Alexander et al. 2008), a behaviour one report correlates with reader interest in the document (Badi et al. 2006). A qualitative study of how readers navigate a leisure-reading periodical may give some insight into this behaviour (Marshall et al. 2005): in that study readers manipulated the periodical to focus their attention, read forward and back for context, and interrupted themselves by turning pages.

Literature summary

Academic books are widely used, and are increasingly used online in some disciplines, and there are differences in the way academic articles and books are found and used.

Online reading, and electronic books in particular, have known usability failings, particularly with respect to document navigation and annotation, two key functions in academic reading. Nonetheless, studies of how users navigate documents are largely limited to articles. While the cues used to assess books and articles for usefulness appear to be similar (though there is a particular emphasis on the table of contents for books), users struggle trying to utilise this information when it is presented in ebooks.

The work presented in this paper is an exploration of natural user navigation behaviour within ebooks, an area not previously studied in either print books or ebooks.

METHODOLOGY

This study used a transaction-log analysis methodology, an ideal methodology for unobtrusively studying user behaviour, though it cannot provide any insight into users' motivations or subjective experiences.

This methodology has been used to study user interactions with a wide range of online systems, including search engines (Jansen et al. 2006), digital libraries (Jones et al. 1998), library catalogues (Lau et al. 2006), and online library resources (Nicholas et al. 2005). While one small study has used transaction logs to gain an overview of ebook usage (though not into how they are used) (Connaway et al. 2005), this method has not been widely applied in the ebook domain. Some likely reasons for this limited uptake are outlined in (Cox 2008), though the largest barrier Cox mentions (incomparable data from different ebook providers) does not apply to this study as it is based on data from a single provider.

To outline how the methodology was applied to ebook use and navigation, it is first necessary to describe the functionalities of the ebook collection; after which more detail will be given about the method as applied here.

The ebook collection

The ebook collection studied is provided by EBL (<http://www.ebllib.com/>), a large ebook provider in Australia. 13172 ebooks are available to library users via our library catalogue. Some of these ebooks are owned by the library and some of them are not but the distinction is transparent to users; all of them are available as part of a patron-driven acquisition model (Hardy et al. 2007). Ebooks in the collection cover a wide range of disciplines, though it should be noted that as the university has a focus on science and technology these books (and their more-frequent users (Christianson 2005; Christianson et al. 2005)) might be slightly over-represented. This collection is intended to be used for academic purposes, and any recreational reading of these books is incidental.

EBL logs a range of anonymized statistics about readers' use, including usage by book, viewing time at the book level in each session, how frequently each book is accessed, and the page numbers viewed in a session in the order in which they are viewed.

Readers navigating from the catalogue to an EBL ebook are delivered to a summary page which includes a cover image for the book and may include a brief abstract or table of contents. From this page they click 'read online', taking them to an 'Intro' page (not collected in the page view statistics). This page usually displays the cover of the book (see Figure 1 overleaf) and also displays a hyperlinked table of contents.

The initial access period for any ebook is termed 'browsing'—an activity that does not require any specific action on the part of the user to express interest and costs

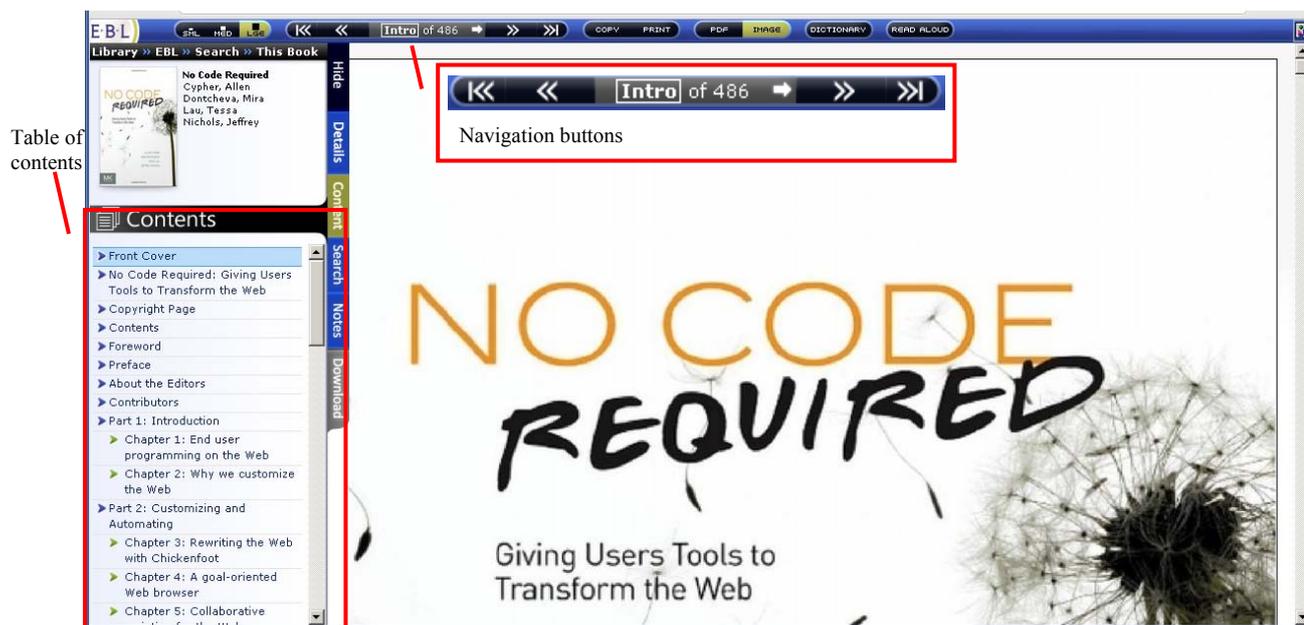


Figure 1: First page of an ebook, showing navigation tools

the library nothing. Readers are required to ‘loan’ an ebook (by clicking a button on a dialog box) to continue reading after ten minutes for if the ebook is owned by the library, or after five minutes if the ebook is paid for on a single-loan basis, or when they print material from any ebook.

From the introductory page users may ‘browse’ within unowned ebooks (those ebooks for which ownership rights have not been purchased) for up to five minutes, or owned ebooks for up to ten. At the end of this time, or when a user attempts to print any pages, the EBL interface forces them to click to create a ‘loan’. The library is charged a fee for the loan if the book is unowned; if it is owned this loan counts against a fixed number of available loans per year. This distinction is transparent to the user. Forcing users to click to create a loan, while specifically used to determine when the library should be charged also means that loaned ebooks are those with which readers engaged sufficiently to expend an (albeit minor) effort to read or use further. Once a book is loaned, readers can print a certain proportion of it, though due to technical considerations with our configuration of EBL even loaned ebooks cannot be downloaded.

Within an EBL ebook, readers can move between pages using page navigation keys and a ‘jump-to’ functionality (see Figure 1). They can also navigate using a traditional right-hand scroll bar. The table of contents (shown in the literature reviewed above to be a vital navigational feature of books) forms a hyperlinked left hand menu. When moving quickly between pages, there is a lag of up to one second before a new page is displayed; only once a page has been displayed is it counted in usage statistics. While this navigation structure is highly likely have usability problems, these problems remain constant between books and thus do not affect results comparing different types of use.

Analysis

The analysis presented in this paper is based on all ebook usage in May, a month that fell entirely during term time at Swinburne.

The first step of this analysis was to calculate usage statistics for ebooks within the Swinburne collection, and compare this to usage statistics from other institutions given in the literature.

The investigation of basic statistics was followed by an exploratory evaluation of document triage practices in this system. The EBL system lends itself to exploring document triage as users engage in a tangible indication of interest when they click to create a loan. Browse data is recorded separately from loan data, however, even for books used in the same session. To get a true picture of document triage requires browse and loan data from the same session to be matched—this was done on the basis of user identification tokens, timestamp and item title. During this process 666 items that had been loaned but not browsed were discovered; these were discarded from the data set as the decision point available in other transactions was not present in this data. Once the data had been processed, loaned and browsed items were compared across a number of features.

It is possible to examine users’ navigation paths through books: whether they ‘flip’ back and forth can be easily determined from examining page number sequences, as can whether they are flicking consecutive pages or turning to more distant parts of the book. For this analysis pages were grouped into ‘chunks’ of text read. If pages viewed consecutively were two or fewer than two pages distant they were considered to be part of the same ‘chunk’. Within chunks readers may move in a single direction, or flip back and forth.

RESULTS

The results of this study are presented in three sections: usage statistics, document triage practice, and navigation within ebooks.

Usage statistics

The test period comprised 12179 sessions, 11488 of which involved access to an ebook. The disparity between the number of sessions and the number of ebook accesses likely comes from the first possible opportunity users have to reject an ebook after locating it in the catalogue: seeing the book in EBL, and choosing not to click 'read online'. The statistics generated by EBL do not specify how many ebooks a reader accessed in a session, so it is impossible to tell whether the proportion of users who successfully accessed an ebook is as high as it seems (94.32%). It is important to note that data cleaning (described in the previous section) discarded 666 sessions, leaving 10954 ebook uses to be considered in this study.

Like the data in (Christianson 2005), usage in our ebook collection fits a power-curve distribution, with most books in the collection (6745 of 13179, or 51.2%) remaining entirely unused, and a very small number attracting a significant amount of use. Figure 2 presents usage of the ebooks in our collection, with the number of books on the y-axis and the number of times they were accessed on the x-axis; note that the y-axis is in a logarithmic scale. The most frequently used book (*Introduction to Theories of Popular Culture*) was accessed 74 times, approximately two-and-a-half times per day. This frequent use suggests that this book may have been used as a textbook or course reading during this timeframe. Another book, *Cultural Theory: The Key Concepts* was the 5th most-used book in the collection, reinforcing this possibility.

Of the books that were used, a range of usage activities were seen, from rejecting an ebook from the introductory page to in one case viewing 794 pages (in a 264 page book, representing reading the book nearly three times in the session) over 509 minutes (an eight and a half hour period). When that session is examined more closely, it is apparent that the reader moved backward and forward through the book a number of times, perhaps reading, taking notes or cross-referencing, echoing some of the activities seen in the literature discussed earlier.

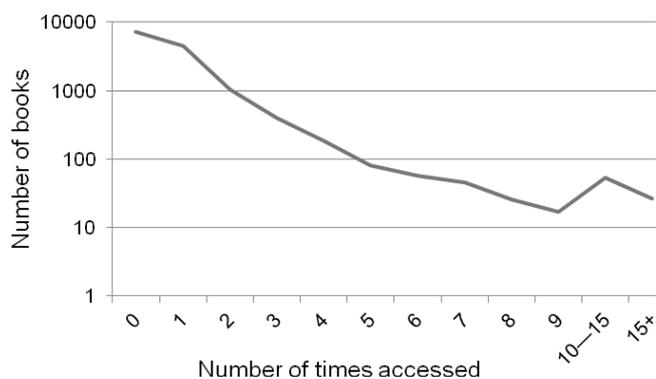


Figure 2: Ebook usage per title

For both the number of page views and the number of minutes browsing, the reading data (i.e. the data about books that were loaned) has long tails: the mean number of minutes per session is 52, median 12 and standard deviation 139.7. Similarly the mean number of pages read is 60.3, median 44, and standard deviation 58.1. While the browse data has some variance in the number of pages read (mean 19.1 pages, median 15, standard deviation 16.6), the ten minute maximum viewing time (before users are forced to loan an item) ensures that the variance in number of minutes spent browsing is limited (mean 1.8, median 1, standard deviation 2.19).

Document triage

'Document triage' refers to the practice of rapidly evaluating documents to determine whether they will meet an information need.

There is a base marker of user interest in the EBL system: the loan. Because users have to take an action to create a loan, they have definitively demonstrated that the ebook in question has further value to them. To assess document triage strategies usage data was compared between loaned books and browsed books.

Users browsed a far greater number of ebooks than they read, a finding that tallies with the number of documents selected as 'interesting' in other studies (and to a certain extent with the practice of search—it would be a rare search indeed that returned only and exactly what a user needed). There are a number of other differences in usage behaviour between browsed and loaned ebooks, including the number of pages read, and time spent reading.

The browsing data reveals 90 books in which users did not get beyond the introductory page (about 1.3% of browsing sessions), a behaviour not seen at all in reading sessions. This suggests that some users make decisions about the usefulness of material based on the cover image and the table of contents shown for every item, similar to the use of metadata users demonstrated with print books in (Stieve et al. 2006). It may be possible to determine which features cause users to reject books so quickly by comparing rejected books with browsed books; possibilities include the length of the book and limited metadata on this introductory page. Such a comparison is, however, outside the scope of the work presented here.

As we would expect, given that browsing is at least partly time limited, users read more pages and spend longer in books where they create a loan than in those they only browse (these differences are both statistically significant to $p < 0.01$). This difference persists, though, even when comparing books browsed or read for less than five minutes (the minimum time after which a reader would be asked to create a loan): readers still view more pages—and spend longer—reading ebooks they go on to loan (again these differences are significant at $p < 0.01$). While this does not provide any insight into how users are determining the usefulness of books, it demonstrates that evaluation takes place early in the reading process.

It is useful to consider *how* users move through books while browsing (as compared to reading). To compare

behaviour chunks' of text read (as described in the previous section) were analysed, again comparing the first five minutes of reading with the first five minutes of browsing. Behaviour in books readers went on to loan was markedly different than in the books they merely browsed. The mean number of pages in a chunk was slightly greater when reading than when browsing (4.07, and 3.69 respectively), though this difference is not statistically significant. The number of chunks was also greater when reading than when browsing (7.51 and 4.84 respectively), a difference that is statistically significant at $p < 0.01$. Finally browsed chunks were more likely to be read in one direction, rather than containing 'flipping' behaviour (which happens in 14.42% of browsed chunks, and 17.45% of read chunks). This difference makes browsing look very similar to 'scanning' as described in (Adler et al 1998) as quick reading for the purpose of decision making, while reading (even in the first five minutes) looks more like the deep reading seen in (Marshall and Bly 2005). This is an interesting finding, and suggests that many of the novel techniques for triaging other kinds of documents (for example the zooming techniques seen in (Buchanan et al. 2008)) are also likely to be useful in triaging ebooks.

Navigation and reading

The data collected by EBL makes it possible to examine for the first time how readers navigate within books, and provides some insight into what reading looks like in an ebook context. This section does not further consider browsing behaviour as it is described in the previous section and is quite distinct from reading behaviour. Instead, this section focuses on user interactions with those texts they with which they actively engaged by clicking to create a loan.

The data collected for this study describes 3859 reading sessions, encompassing viewing of some 232514 pages over a period of 3405 hours (or approximately 56 days). This data was examined to determine how sequentially users read and how they moved through books more generally, what their reading looked like under different circumstances, and what cues they were using to navigate within books.

Always going forward?

The same chunks used to compare reading and browsing behaviour in document triage were used to determine how sequentially readers moved through content while reading. Chunks were on average four pages long, and the median number of pages flipped between chunks was 5, however there was a long tail of distance flipping: the mean number of pages flipped was 38.0. The largest number of pages flipped past in a single action was 1163—an improbable feat while reading a physical book, given their general heft. The total number of pages read is positively correlated with the number of chunks (Pearson coefficient=0.79), suggesting that flipping behaviour remains constant regardless of the number of pages read.

While only one reader began reading anywhere other than the first page (beginning instead on page 2), the reading behaviour seen in these ebooks is far from the sequential

behaviour we could expect of (for example) novel reading. 17% of all chunks contained internal forward and backward flipping, and only 3% of sessions did not include at least one move backward between chunks. Users flipped forward just over half the time (52.5%), however when we consider the larger jumps—those greater than the mean of 38 pages—this sinks to 40.5% (a difference significant at $p < 0.01$). This dramatic difference suggests a stepwise movement forward through books with larger jumps back.

When examined at the macro level, the page flow data reveals four major paths through the pages viewed in a session: moving consistently through the pages viewed, moving through a large number of pages and then working backwards, making a significant jump forward through a number of pages then staying in the same area of the book and moving stepwise through examined pages with big jumps interspersed with proximal reading. These paths, represented in Figure 3, are similar to those seen in (Loizides et al. 2009), suggesting that either this pattern is used for more than just triage for scholarly material, or that ebooks are constantly being triaged.

While strictly sequential reading in ebooks is hardly ever seen in this data set, the trend (with the exception of the large jumps back) is generally to begin near the beginning of a book and work forwards.

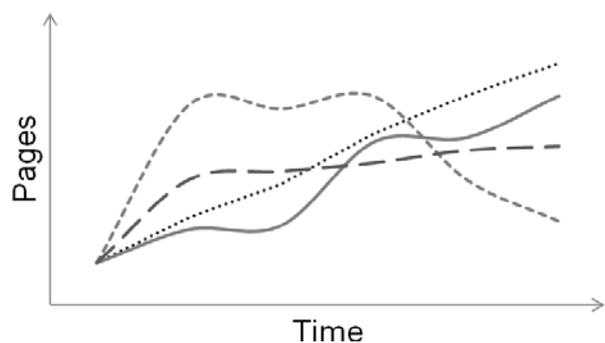
Reading activities

It is unfortunate that EBL does not log viewing time at the page level; this would provide considerable insight into what users are do on specific pages. Time per book is logged, however, and this can be used to calculate an approximation of view time per page. View time per page in during reading ranges from as little as one second (suggesting pages loading as users scrolled past) to 67.1 minutes. There is no correlation between time and number of pages viewed, or between time and number of chunks suggesting that for any given session readers may be closely engaged with a single page, or scrolling through a large number of pages.

Navigational cues

A manual analysis was necessary to investigate whether ebook readers use the same navigational cues as those reading other kinds of scholarly material.

It is evident from the literature described earlier that, at least in other scholarly material, readers change direction and slow down near textual features such as section headings, titles, and emphasized text. To test whether this behaviour also exists in ebook, reading six groups of 80 pages were examined. The different groups were single



page chunks, maximum and minimum page numbers within chunks, first and last pages of chunks, and finally, for comparison with pages where users were not changing direction, middle pages of chunks.

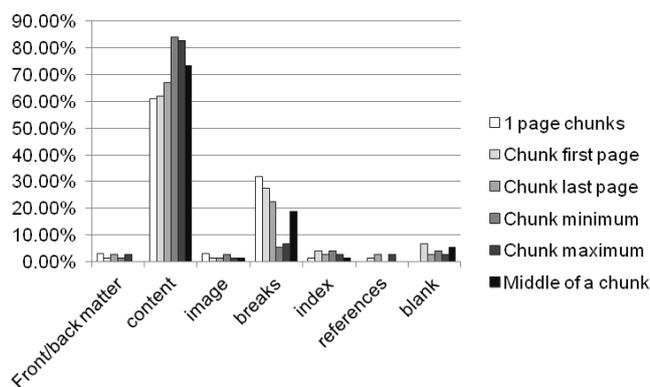


Figure 4: Navigation features in text and change of direction

Each sample page was examined in the relevant ebook to determine whether the page contained a section or chapter break, or was an image, front or back matter, an index page, or plain content.

Plain content dominates these figures, likely because it dominates book content generally. However there were some differences in content according between groups; Figure 4 above shows the distribution of content type for different chunk positions. While the various chunk positions appear to broadly similar in page-type distribution the differences between them are significant ($\chi^2=45.092$, $df=30$, $p=0.038$). Single page chunks and the first pages of chunks include content breaks (for example chapter headings) more frequently than other kinds of content; suggesting users are using the table of contents to navigate within ebooks. Similarly, last pages of chunks also frequently include section breaks, though not to the same degree as first pages and single page chunks (this may be because last pages are also likely to be the page *before* a section break). These findings suggest that readers use the navigation features in ebooks as in similar ways to those in other kinds of documents.

Summary of results

Ebooks in this collection were well used, with total usage time amounting to slightly more than 56 days (the equivalent of two people using the ebooks nearly constantly for the test period of one month). Use comprised 11488 sessions; ebooks in this study were used in the same power curve distribution as resources in other studies, for example (Christianson 2005).

The data collected allows a clear distinction to be drawn between browsing and reading sessions; the latter requires readers to click a button to confirm their interest. There were notable differences in behaviour between browsing and reading, including viewing fewer pages and reading more sequentially when browsing.

A close examination of reading behaviour showed that reading is not sequential in academic ebooks; in fact ebook reading patterns appear more like those seen in

journal articles or other scholarly materials than what we could expect from fiction reading. Another similarity between ebook and journal article reading is the use of document structure to provide navigational cues; in ebooks chapter headings and other breaks are likely to trigger a change in page location or reading direction.

DISCUSSION

As described earlier, ebooks and ebook reading devices are not without their usability problems (see for example (Malama et al. 2004; Rowlands et al. 2007; Shelburne 2009; Berg et al. 2010; Li et al. 2011)). Nonetheless ebook use in academic and leisure reading has increased dramatically in the past five years (Hamblen 2011; Li et al. 2011), a trend that seems likely to continue. This increase is likely partly because ebooks afford convenient access to material that readers may not otherwise elect to use (McKay 2011) or be able to access (Hardy et al. 2007). Similarly, ebook technology can make library collections more responsive to user needs for example by facilitating patron-driven book purchase (Hardy et al. 2007). Despite these advantages, it is important to provide the best possible ebook reading experience to users. In order to do this we need to understand what users do with and want from ebooks.

In this study we saw a range of behaviours that seem broadly similar to reading of non-book academic and scholarly materials: readers appeared to triage documents by scanning (O'Hara et al. 1997; Marshall et al. 2005), move back and forth through a document when reading more closely (O'Hara et al. 1997; Marshall et al. 2005) and use document structure to facilitate reading (Bishop 1998; Loizides et al. 2009). These similarities are interesting in light of the work showing that users of academic libraries perceive and search for books differently from other materials (McKay et al. 2010; McKay et al. 2011), and that different disciplinary cultures use books at different rates (Talja et al. 2003; George et al. 2006).

It is unclear from this study whether the similarities between ebook and scholarly article use are due to electronic format, though some article usage behaviour appears to be format-independent (Loizides et al. 2009). It is possible that these similarities are not format-related, but instead reflect a consistency in scholarly and academic information practices that belie the perceptions and information seeking behaviour of readers. Further investigation of readers' use of academic print books could shed light on this issue.

Whatever the reason for the similarities between ebook reading and reading of other scholarly materials, the behaviour seen in this study (flipping through books, moving backward and forward, and using document structure for navigation) are poorly supported by computer screens and dedicated ebook reading technologies alike (Pearson et al. 2010; MacFayden 2011; Thayer et al. 2011). Given this disparity, there is considerable scope to develop more reader-friendly ebook reading technologies for academic purposes.

CONCLUSIONS

Ebook technology affords a unique opportunity to HCI practitioners: the ability to examine naturalistic book reading behaviour, an activity which is difficult or even 'creepy' to study in other ways (Marshall 2010). This paper presents an exploratory log analysis of just such behaviour; to the author's knowledge it is the first study of book use irrespective of format.

This work examined ebook usage, document triage, and in-book examination while reading. The ebook usage statistics that underpin this investigation are broadly similar to those from other institutions, suggesting the behaviour seen in this study is broadly representative of ebook use.

Because users have to click to create a book loan in the EBL system, it is possible to compare browsing behaviour (where the user did not have enough interest in the document to create a loan) with reading behaviour, (where the reader confirms their interest in a document with a click). Reading behaviour is significantly different from browsing behaviour, even before the point at which users would have been required to create a loan: flipping back and forth was more common during reading, and a greater number of pages were read.

Reader navigation through ebooks was also investigated. Reading was discovered to be largely non-sequential at the micro level though readers for the most part moved forward through ebooks at the macro level. Macro-level reading patterns were fairly similar to those described in the literature for journal articles, as was use of cues such as headings and breaks for navigation.

This study is our first indication that academic ebook reading behaviour is similar to reading behaviour in other scholarly documents. These similarities render many ebook reading technologies inappropriate for academic use as they do not support flipping through books and they often reduce navigation cues.

This study sheds considerable light on ebook reading behaviour, but it is unclear whether these findings are format-specific, or whether academic print books are read in similar ways. This question bears further investigation, as understanding reading practices in print books may further enhance our ability to design effective ebook reading technologies. Similarly, users' subjective experience of reading and the reasons for the behaviours seen in this study warrant further investigation: log analysis cannot answer these questions, so alternative methods need to be applied to this question.

ACKNOWLEDGEMENTS

Thanks to Mike Scott and Rebecca Parker for their comments on this document. Thanks to Tony Davies for getting me to understand EBL usage data.

REFERENCES

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. Proc.

- CHI '98, ACM Press/Addison-Wesley Publishing Co. (1998), 241-248.
- Alexander, J. and Cockburn, A. An empirical characterisation of electronic document navigation. Proc. Graphics Interface 2008, Canadian Information Processing Society. (2008).
- Badi, R., Bae, S., Moore, J. M., Meintanis, K., Zacchi, A., Hsieh, H., Shipman, F. and Marshall, C. C. Recognizing user interest and document value from reading and organizing activities in document triage. Proc. IUI '06, ACM. (2006).
- 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, (2006) 1-27.
- Bailey, T. P. Electronic Book Usage at a Master's Level I University: A Longitudinal Study. The Journal of Academic Librarianship 32, 1, (2006) 52-59.
- 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, (2010) 518-525.
- Bishop, A. P. Digital libraries and knowledge disaggregation: the use of journal article components. Proc. DL '98, ACM. (1998), 29-39.
- Blustein, J., Rowe, D. and Graff, A.-B. Making Sense in the Margins: A field study of annotation. Proc. TPD 11, Springer. (2011), to appear.
- Buchanan, G. and McKay, D. In the Bookshop: Examining Popular Search Strategies. Proc. JCDL 2011, ACM. (2011).
- Buchanan, G. and Owen, T. Improving skim reading for document triage. Proc. IIX '08, ACM. (2008), 83-88.
- Christianson, M. Patterns of use of electronic books. Library Collections, Acquisitions, and Technical Services 29, 4, (2005) 351-363.
- Christianson, M. and Aucoin, M. Electronic or print books: Which are used? Library Collections, Acquisitions, and Technical Services 29, 1, (2005) 71-81.
- Connaway, L. S. and Snyder, C. Transaction Log Analyses of Electronic Book (E-book) Usage. Against the Grain 17, 1, (2005) 85-89.
- Cox, J. Making Sense of E-book Usage Data. The Acquisitions Librarian 19, 3, (2008) 193 - 212.
- Dillon, A. Reading from paper versus screens: A critical review of the empirical literature. Ergonomics 35, 10, (1992) 1297-1326.
- George, C., Bright, A., Hurlbert, T., Linke, E. C., St. Clair, G. and Stein, J. Scholarly use of information: Graduate students' information seeking behavior. Information Research 11, 4, (2006).
- Greifeneder, E. The Impact of Distraction in Natural Environments on User Experience Research Proc. TPD 11, Springer. (2011), to appear.

- Hamblen. Amazon: E-books now outsell print books, Computerworld (2011). Available at: http://www.computerworld.com/s/article/9216869/Amazon_E_books_now_outsell_print_books
- Hardy, G. and Davies, T. Letting the patrons choose: using EBL as a method for unmediated acquisition of ebook materials. Proc. Information Online 2007, ALIA. (2007).
- Hernon, P., Hopper, R., Leach, M. R., Saunders, L. L. and Zhang, J. E-book Use by Students: Undergraduates in Economics, Literature, and Nursing. The Journal of Academic Librarianship 33, 1, (2007) 3-13.
- Jansen, B. J. and Spink, A. How are we searching the World Wide Web? A comparison of nine search engine transaction logs. Information Processing & Management 42, 1, (2006) 248-263.
- Jones, S., Cunningham, S. J. and McNab, R. An Analysis of Usage of a Digital Library. Proc. ECDL'98, Springer. (1998), 261-277.
- 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, (2006) 1316-1329.
- Li, C., Poe, F., Potter, M., Quigley, B. and Wilson, J. UC Libraries Academic e-Book Usage Survey. University of California, Springer, California Digital Libraries (2011).
- Liesaputra, V. and Ian, W. H. Seeking information in realistic books: a user study. Proc. JCDL '08, ACM. (2008), 29-38.
- Littman, J. and Connaway, L. S. A Circulation Analysis of Print Books and E-Books in an Academic Research Library. Library Resources & Technical Services 48, (2004) 256-262.
- Loizides, F. Investigating Document Triage on Paper and Electronic Media. Proc. ECDL 2007, Springer Berlin / Heidelberg. (2007), 416-427.
- Loizides, F. and Buchanan, G. An Empirical Study of User Navigation during Document Triage. Proc. ECDL 2009, Springer Berlin / Heidelberg. (2009), 138-149.
- MacFayden, H. The Reader's Devices: The affordances of ebook readers. Dalhousie Journal of Interdisciplinary Management 7, 1, (2011).
- Malama, C., Landoni, M. and Wilson, R. Fiction Electronic Books: A Usability Study. Proc. ECDL '04, Springer Berlin / Heidelberg. (2004), 69-79.
- Marshall, C. C. Reading and Writing the Electronic Book. Chapel Hill, NC USA, Morgan & Claypool (2010).
- Marshall, C. C. and Bly, S. Turning the page on navigation. Proc. JCDL'05, ACM. (2005), 225-234.
- McKay, D. Gotta keep 'em separated: Why the single search box may not be right for libraries. Proc. CHINZ '11, ACM. (2011), to appear.
- McKay, D. and Buchanan, G. One of these things is not like the others: how users search different information resources. Proc. TPDFL 11. (2011), to appear.
- McKay, D. and Conyers, B. Where the streets have no name: how library users get lost in the stacks. Proc. 11th Annual ACM SIGCHI NZ Conference on Computer-Human Interaction (CHINZ 2010), ACM. (2010), in press.
- Nicholas, D., Huntington, P. and Watkinson, D. Scholarly Journal Usage: The Results of Deep Log Analysis. Journal of Documentation 61, 2, (2005) 248-280.
- O'Hara, K. and Sellen, A. A comparison of reading paper and on-line documents. Proc. CHI '97, ACM. (1997).
- Ooi, K. How Adult Fiction Readers Select Fiction Books in Public Libraries: A Study of Information Seeking in Context. School of Information Management. Victoria University of Wellington (2008).
- Pearson, J., Buchanan, G. and Thimbleby, H. HCI design principles for ereaders. Proc. Booksonline '10, ACM. (2010).
- Rettig, J., Bandelin, J. and Park, L. E-books: what have we learned? Where are we going? Proc. ACRL National Conference, ACRL. (2003).
- Rowlands, I., Nicholas, D., Jamali, H. R. and Huntington, P. What do faculty and students really think about e-books? Aslib Proceedings 59, 6, (2007) 489-511.
- Shelburne, W. A. E-book usage in an academic library: User attitudes and behaviors. Library Collections, Acquisitions, and Technical Services 33, 2-3, (2009) 59-72.
- Stelmaszewska, H. and Blandford, A. From physical to digital: a case study of computer scientists' behaviour in physical libraries. International Journal on Digital Libraries 4, 2, (2004) 82-92.
- 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, (2006) 599-608.
- Talja, S. and Maula, H. Reasons for the use and non-use of electronic journals and databases: A domain analytic study in four scholarly disciplines. Journal of Documentation 59, 6, (2003) 673-691.
- Tenopir, C., King, D. W., Edwards, S. and Wu, L. Electronic journals and changes in scholarly article seeking and reading patterns. Aslib Proceedings 61, 1, (2009) 5-32.
- Thayer, A., Lee, C. P., Hwang, L. H., Sales, H., Sen, P. and Dalal, N. The imposition and superimposition of digital reading technology: the academic potential of e-readers. Proc. CHI 2011, ACM. (2011).