Judging a book by its cover: interface elements that affect reader selection of ebooks

Dana McKay
Library, Swinburne Institute, Swinburne University of Technology
Hawthorn, Melbourne, Australia
dmckay@swin.edu.au

George Buchanan
Centre for HCI Design, City University
Northampton Square, London, UK
george.buchanan.1@city.ac.uk

Nicholas Vanderschantz, Claire Timpany, Sally Jo Cunningham, Annika Hinze
Department of Computer Science, University of Waikato
Hamilton, New Zealand
{vtwoz, ctimpany, sallyjo, hinze}@waikato.ac.nz

ABSTRACT
Digital library research has demonstrated the impact of content presentation on both search and reading behaviours. In this paper, we scrutinise the influence of ebook presentation on user behaviour, focussing on document thumbnails and the first page view. We demonstrate that flaws in presentation increase the volume of short time-span reading, and reduce the likelihood of long-span reading when compared to other documents. This reflects other patterns of information seeking behaviour that demonstrate increased short-term reading when information content is uncertain, and suggests an ineffective use of reader time on less useful content.

Author Keywords
Ebooks, information seeking, book selection, search interfaces

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
It is a core principle of HCI that interface, interface elements and interface errors affect user behaviour (Dix et al. 2004). This is demonstrably true in information seeking interfaces, as it is in other types of interface (see for example (Mahoui et al. 2000)). This principle also holds true in physical book retail and library environments: readers are more likely to interact with display books than books shelved with only their spines visible, for example (Ooi 2008; Buchanan et al. 2011; Hinze et al. 2012).

Increasingly, though, book interaction is occurring away from these physical environments: in 2011 ebook sales eclipsed print book sales on Amazon for the first time (Hamblen 2011). In some academic disciplines ebook use in libraries surpassed print circulation as early as 2004 (Littman et al. 2004). Ebook use across disciplines has increased with each successive large study; compare, for example, Rowlands' 2007 (Rowlands et al. 2007) work, in which 44% of respondents had used an ebook, with Li's work in 2011 (Li et al. 2011) where 58% of respondents reported ebook use.

Despite these increases in ebook usage, it is clear that for both academic and recreational purposes ebooks still have significant usability issues. Readers are frustrated and confused by DRM restrictions on access (Shelburne 2009; Marshall 2010), bemoan the lack of good tools for annotation (Pearson et al. 2010; Thayer et al. 2011), and struggle to navigate within both fiction (Malama et al. 2004) and academic texts (Liesaputra et al. 2008; Berg et al. 2010). Given these problems it is reasonable to assume that readers also struggle with ebook technologies for selecting ebooks and determining their usefulness.

Ebook selection practices have significant financial implications for booksellers and libraries alike. Booksellers are rewarded when readers elect to purchase a book, but as in other types of ecommerce, must make books appealing enough to convince readers to part with their money (Holt 2010; Hamblen 2011). Libraries, conversely, pay for ebook access using different models (see for example (Hardy et al. 2007; Cummings 2011)), : they may pay when readers view any part of an ebook, or they may pay only after a certain time has elapsed or proportion of an ebook has been used. In any payment model, though, beyond a certain point of usage libraries will pay for ebook use, and thus it is in their interest to ensure that this use is valuable to their patrons, rather than patrons merely browsing and rejecting a string of books.

In light of all this, it is perhaps surprising that so little work exists on ebook selection practices, and the interface elements that guide these practices. While this dearth of work does reflect the lack of information about book selection practices more generally (Rowlands, Nicholas et al. 2007), ebook technologies afford unique opportunities for study of reader behaviour. Transaction log analysis is a long-standing technique used to understand information seeker behaviour in online library catalogues, (Cousins 1992; Cooper 2001) digital libraries (Jones et al. 1998; Mahoui et al. 2001) and even search engines (Jansen et al. 2006): the advent of ebooks allows us to bring this technique to bear on parts of the reading process that have previously been very difficult to study (Marshall 2010).

OZCHI'12, November 26–30, 2012, Melbourne, Victoria, Australia.

This paper presents the results of a study using ebook publisher transaction logs from an academic library collection to determine which user interface elements affect reader selection behaviour. We present our work in five sections: first we discuss the background and motivation for this work, next we describe our study methodology, thirdly we present our analysis of our results, fourthly, we then discuss our findings in light of other work in this area, and finally we draw conclusions and suggest avenues for future work.

BACKGROUND AND MOTIVATION

That system interfaces influence human behaviour is a central tenet of HCI. This principle is most heavily relied upon in safety-critical systems (see for example (Casey 1998; Salvucci 2001)). Work on information systems, however, has also demonstrated that user information behaviour is affected by interface design: for example Jones et al showed that users requested more search results for unranked Boolean queries than ranked results in 1998 (Jones, Cunningham et al. 1998). More recently, and reflecting a larger change in interface, Ballard et al demonstrate that users are 15-20 times more likely to refine their search queries in a new generation catalogue than a traditional library catalogue (Ballard et al. 2011).

The use of ebooks is increasing, both in the academic (Littman and Connaway 2004) and recreational (Holt 2010) spheres. While the business models libraries use to make these books available to their readers may differ, in all models libraries must pay for patrons' access to ebooks in some manner. In light of this, it is important that readers are best able to select books that interest them in a manner that allows libraries to only pay for books that are actually useful to readers. In addition to the financial cost imposed on libraries, there is a user cost to information seeking systems that do not allow users to rapidly and effectively locate the most useful results: Information seekers demonstrably turn to systems that they know return poorer results in exchange for convenience (Ponsford et al. 2007), avoid seeking resources that are difficult to access (Connaway et al. 2011), satisfice (Agosto 2002), or just give up (Nordlie 1999). It is perhaps surprising, then, that so little attention has been paid to the effect of interface (and individual interface elements) on ebook selection.

Selection among available information resources generally has been relatively well-studied: it is clear that users are more likely to investigate items that fall higher in a list of search results (Jansen and Spink 2006), for example, and that inconvenient items (in terms of convenience defined by the user) are less likely to be used (Connaway, Dickey et al. 2011).

Rowlands noted in 2007 (Rowlands, Nicholas et al. 2007) that the process of selecting a useful book from the available options is a surprisingly under-studied part of the book selection process for *all* books, not just ebooks. This statement remains largely true, with some exceptions: Reutzel and Gali (Reutzel et al. 1998) studied children selecting fiction books in a physical library, and noticed they were influenced by shelf position and cover, rather than content. Moore's work (Moore 1995) demonstrated the influence of shelf position on children's selection practices in fiction books, and Borgman (Borgman et al. 1995) demonstrated the same bias in digital libraries. Among adults, cover clearly does have an influence in both bookshops (Buchanan and McKay 2011) and academic libraries (Stelmaszewska et al. 2004; Hinze, McKay et al. 2012). It is not the only influence, however, Stieve showed in 2006 (Stieve et al. 2006) that when choosing between two similar books university students relied heavily on the table of contents; a behaviour that was also demonstrated "in the wild" in our own earlier work on both physical and digital academic libraries (Hinze, McKay et al 2012; McKay, Hinze et al. 2012). Finally, both Stelmaszewska (Stelmaszewska and Blandford 2004) and our own earlier work (Hinze, McKay et al 2012; McKay, Hinze et al. 2012) show that book content has an impact on decision making in both physical and digital book libraries. . A striking aspect of this decision process is how quickly many selections are made, though how users assess content rapidly remains an open research question. It seems, then, that in rapid decision making in print book selection cover image and table of contents play a significant role, however how these artefacts affect the decision making process remains unclear.

This dearth of book selection research is also surprising in light of readers' propensities for giving up their search: in the online space there are no intermediaries—librarians (Crabtree et al. 1997; Nordlie 1999) or shop assistants (Buchanan and McKay 2011)—to guide readers to the right resources. These difficulties with book search and selection likely result in lost sales (in online bookshops) or unwarranted paid use (in libraries, where readers perceive the books as "free"(Hernon et al. 2007)) over and above the frustration experienced by users.

In addition to the dearth of material on book selection, there is little material available on reading, even for ebooks, perhaps because most methods of studying reading are so overt and obtrusive as to result in a "creepy" experience for the reader (Marshall 2010). What little literature there is shows that content presentation influences reading behaviour. At the micro level, readers demonstrably pay attention to headings and images, both in print (Marshall et al. 2005) and in electronic form (Loizides 2007). More broadly, readers find navigation easier and enjoy the reading experience more in ebooks that are presented in a "realistic book" style over more conventional ebook presentation methods (Liesaputra and Witten 2008). At a macro level, readers find dedicated ebook readers a more pleasant reading experience than ebooks presented on a computer screen (Li, Poe et al. 2011), but readers are still far from happy with the ebook reader experience: in a recent study over 60% of college students accessed their readings in print rather than reading them on the Amazon Kindles™ given to them by the university (Thayer, Lee et al. 2011). It is evident, then, that content presentation affects not only book selection, but reading patterns as well.

The motivation for our study in particular comes from the realisation that 13-15% of users of the ebook collection in

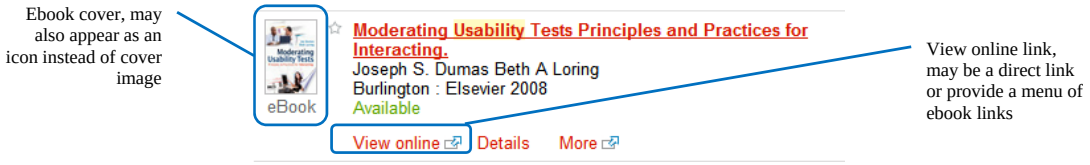

Figure 1: Ebook catalogue entry

this academic library make a snap decision about the usefulness of ebooks they viewed (choosing to investigate further, print or copy material, or rejecting the book and leaving the system) within the first three pages that they sighted (McKay 2011). This decision comes quickly within the ebook reading interface despite readers having to navigate away from search results through a series of interfaces to reach these pages. This rapid decision making is also surprising given that the pages readers view are near-universally the "front matter" that readers said was not useful and impeded the usability of ebooks in another study (Malama, Landoni et al. 2004), and which contain information that even the most savvy of book shoppers appear not to understand (Buchanan and McKay 2011). It seems likely, then, that readers' decision making was influenced by the interface elements they saw as part of the book selection process—elements including search results, book cover, blurb, and table of contents. We set out to investigate how these elements affected decision making by comparing the elements seen in books users rejected quickly to those in books users rapidly decided to interact with further.

STUDY DESGIN AND METHODOLOGY

As mentioned in the previous section, 13-15% of readers of ebooks in Swinburne Library make a decision to create a loan or abandon a book within three pages. To determine which user interface elements may have affected readers' decision making process, we used a grounded theory analysis (Glaser et al. 1967). This section is broken down into an overview of the technology underlying ebook loans in this study, a description of how we created our classification scheme and the classification process, and an outline of how our sample data were selected.

Ebook technology

During the study period (September 2011), Swinburne Library made approximately 19,400 ebooks available to readers via the library catalogue from EBL (an educational ebook publisher). Not all of these ebooks were owned by the library; a number of books were available on a per-demand basis, and purchased only after three users had accessed them via the catalogue (a model known as patron-driven-demand (Hardy and Davies 2007)). This distinction is largely transparent to the user: the only difference from their point of view is limited metadata in the catalogue for books which have not yet been purchased, and a shorter browsing time before the system asks them to click to create a loan (5 minutes as opposed to 10 minutes).

For the purposes of this study, accessing an ebook in EBL from a reader's point of view consists of three steps:

1. finding the book in the catalogue (see catalogue entry in Figure 1),
2. clicking through to an information page about the ebook within EBL, the ebook publisher (see splash-page in Figure 2), and then
3. clicking through to the ebook interface (see Figure 3 overleaf).

Ebooks may be used in one of two ways: browsing (free for the library, and requiring no specific action on the part of the user) and taking out a loan (incurs a cost to the library—though this is transparent to the user—and which requires the user to click to confirm their continued interest in the book).

There are three conditions under which users must create a loan to continue reading an ebook: if they elect to print some of its content, if they elect to copy some of its content, or if they wish to continue to use it beyond a fixed free browsing period. These conditions automatically generate a pop-up box requesting that the user create a loan to continue, thus the creation of a loan is, while a result of user behaviour, not a user-generated action.

Sampling

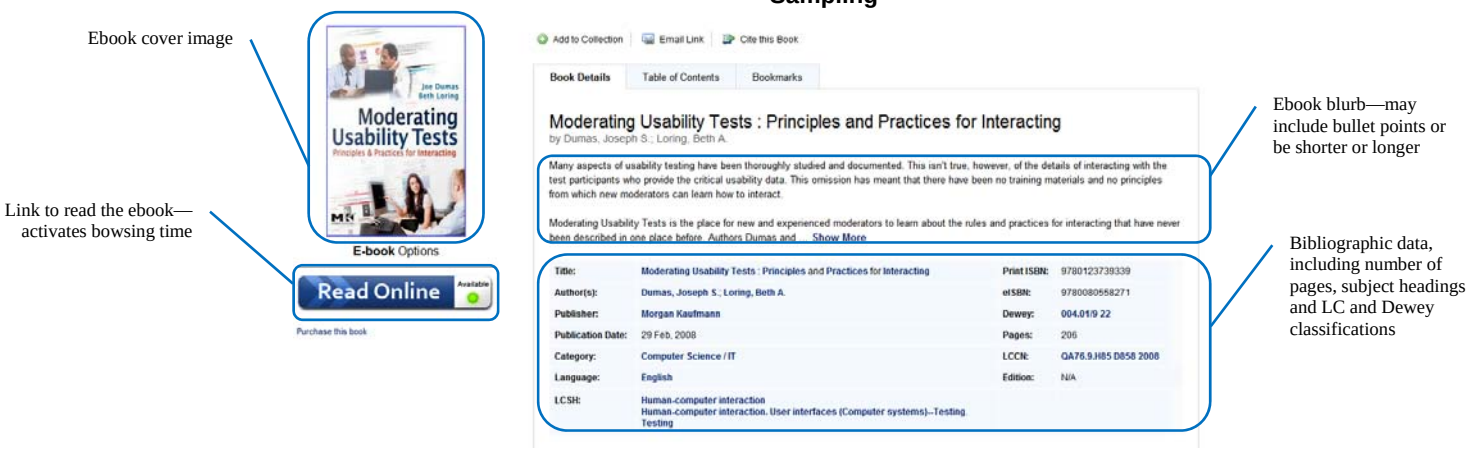

Figure 2: Ebook splash page including blurb, bibliographic data and cover

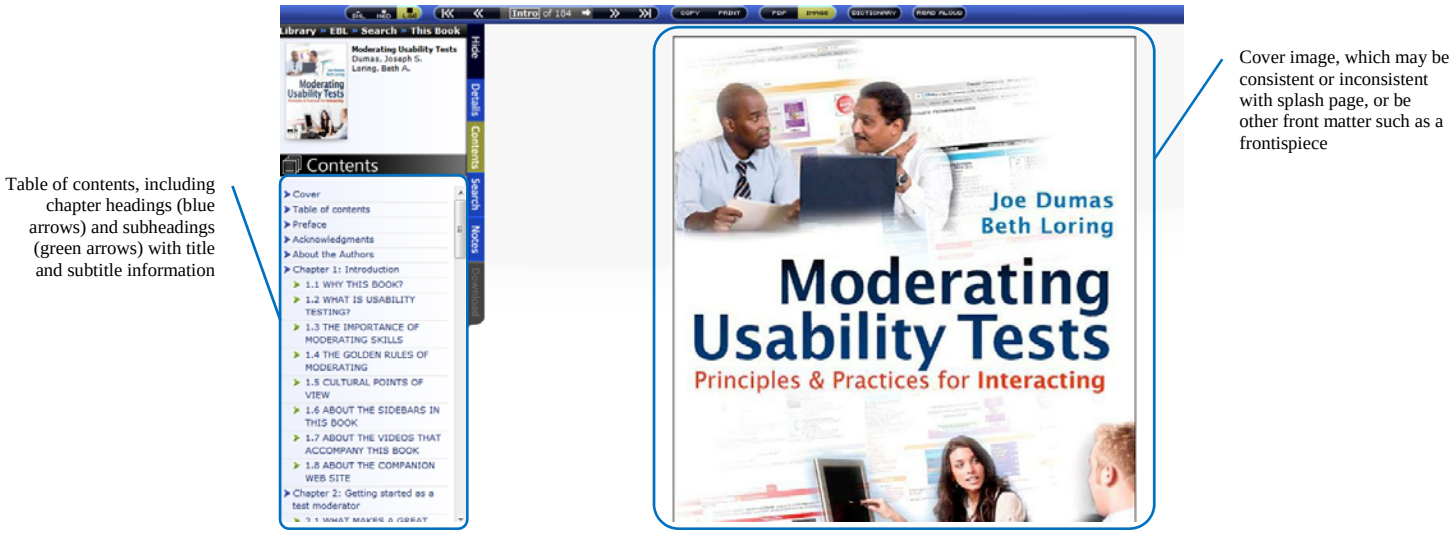

Figure 3: First page of an ebook in the reading interface

The sample period for the study was September 2011, term time for all students at Swinburne University of Technology, a small, research-active dual sector tertiary institution in Australia. During this period, readers accessed 9506 ebooks, going on to create loans in 3225 cases. 914 readers (16%) of those who only browsed and 422 readers who borrowed a book (13%) made their decision within the first three pages of interaction with the book. This discrepancy between browsing and reading is to be expected, as there is no active way to create a loan; those who borrowed a book after the first three pages most likely triggered a loan by copying or printing part of the book.

To compare how the interface elements readers saw affected their decision-making processes, we had to define patterns of decision-making behaviour. The first and most obvious division is between those who browsed books and those who borrowed them: browsing only three pages represents a fairly rapid rejection of a book, while borrowing a book after three pages represents a quick decision that the book is interesting. Rejecting a book quickly is a simple behaviour, and cannot be further classified; our first sample data set comprised catalogue data, ebook splash page, and the first page in the ebook reading interface for 50 unique books that were rejected within three pages and before creating a loan. Choosing rapidly to interact further with a book (by creating a loan) represents a range of behaviours that can only be discovered by investigating after-loan behaviour: when we looked at this behaviour we found four different patterns, and created a sample set of 20 unique books for each of these behaviours:

- **Very short assessment** those who engaged with between 8 and 20 pages in total over the course of 3 or fewer minutes. Readers in this group borrowed books, then either printed or copied content very rapidly before leaving the EBL interface or abandoned the book. Population set size 33 book loans.
- **Short assessment:** Those who read between 8 and 19 pages over the course of their interaction over more than three minutes, but without further restrictions on time. Population of 49 book loans.
- **Medium assessment:** those who read between 20 and 40 pages over longer than 8 minutes. Population set size 64 loans.
- **Long book use:** those who read between 40 and 80 pages over longer than 13 minutes. Population size 68 loans.

Time was only used as a lower limit parameter because it correlates poorly with the number of pages read more generally (this low correlation was also seen in our earlier work on ebooks (McKay 2011)). Many of the books that underwent a very short assessment in number of pages had read times over an hour long; we suspect this reflects readers moving on to some other resource without navigating away from the ebook, for example in a different browser tab. Conversely a number of loans involved a large number of pages accessed in a short space of time, likely due to scrolling. The intervals of page numbers accessed were determined on the basis of population distribution, with the exception of the very short assessment group, which was formed on the basis of our earlier work which demonstrates a large number of users making a decision in three or fewer minutes (McKay, Hinze et al. 2012).

Classification

We used grounded theory analysis (Glaser et al. 1967) to create a classification scheme for the interface elements seen by the user in the process of viewing an ebook. We examined elements within the catalogue entry (see Figure 1), within the ebook landing page (see Figure 2) and the first page of the book visible in the ebook display (see Figure 3). We also examined some elements of the transitions between interfaces

Two coders who are experienced in designing and analysing information interfaces and with some

experience in publishing and digital libraries determined which elements of the ebook selection process would be classified. These coders examined 10 books each and identified elements of the book presentation interface that they believed may affect user selection of books. After 5 books no new classes of difference between books were seen, so the classification scheme was considered complete. Elements coded within the classification scheme included:

- **Within the catalogue entry (see Figure 1):**
  - Catalogue icon: is this a picture (as shown in Figure 1) or an icon
  - Access to alternative formats: was there a physical copy of the book
  - View online: did the view online link (see in Figure 1) link directly to the ebook splash page, or did it present readers with a list of options
  - First result: in a title search, was the ebook the first result returned
- **From the ebook landing page (see Figure 2):**
  - Bibliographic data shown in the ebook splash page, including number of pages, Dewey decimal number(s), Library of Congress classifications and number of Library of Congress classifications.
  - Whether the cover matched the image in the catalogue entry
  - Details about the cover graphics, including field type, image type if present (art, photographic, abstract)
  - Details of cover text, including number of fonts, orientation, and superimposition over images
  - Length and format of the book blurb
- **From the first page of the ebook within the ebook interface (see Figure 3):**
  - Number of pages
  - Type of page for page one (cover image, frontispiece, or other first page)
  - Consistency of the ebook cover with the cover in the ebook landing page
  - Details about left hand table of contents headings, such as whether they included descriptive text or chapter numbers, typical heading length and whether they were presented in all capital letters
  - Details about left hand table of contents chapter numbering
  - Length of left-hand table of contents
  - Whether the left hand table of contents was hierarchical, and whether this hierarchy had been inadvertently flattened.

Once the classification scheme was created, inter-coder reliability was established over a sample of ten ebooks—an initial sample of five was coded, and any differences of opinion were discussed. Five more books were then coded and checked for consistency, which was achieved. The remainder of the loans in our sample were classified by only one of the two researchers involved in coding.

ANALYSIS

Our analysis followed a two-stage process. As described in the previous section, first we coded an initial sample of 20 ebooks from each set to obtain an initial picture of which features might potentially influence user behaviour (based on our log data). This resulted in a shortlist of features that we then coded across a larger sample, to further test our initial hypotheses and to arrive at more reliable data.

The initial coding followed the structure of the interaction sequence of reading and retrieval reported in 'Classification' above.

Many of the books' features were easily dispensed with as demonstrating any potential impact on users' information behaviour. This included much of the basic bibliographic data, including number of pages and the number of Library of Congress Subject Headings. Cover features such as content (photographs or drawings), typographic presentation and general formatting also appeared unlikely to have strong effects. We do not argue that there are no such effects, but it does appear that their impact in this context is at most of minimal power.

	Corrupt cover	Cover then frontispiece	Minor cover inconsistency	Major cover inconsistency	Total cover errors
Not loaned	0.28	0.14	0.20	0.10	0.72
Very short reading	0.29	0.29	0.26	0.19	1.03
Short reading	0.16	0.21	0.16	0.26	0.79
Medium reading	0.10	0.25	0.13	0.10	0.58
Long reading	0.08	0.28	0.00	0.00	0.35

Table 1: cover error rates for different book sets

Where cover data—and some other features—did appear to have an impact on user behaviour was where this data was incomplete or inconsistent. To confirm or deny this hypothesis, we re-validated the consistency of our coding, and extended our samples either to exhaustion, or to fifty items (whichever was the smaller). Full coding was still carried out, but hypothesis testing was focused on those features that had been highlighted in the first round of classification. To ensure rigour we tested both classes (such as cover or table of contents) as a whole, the individual features that appeared to have an impact on behaviour. If behaviour patterns did exist they would first be validated by a class-level analysis and then at the feature level with appropriate checks for over-testing that are appropriate in such cases.

As the comparison of the different sets is a case of proportional test, we applied the Chi-squared test to determine statistical significance, corroborated by more advanced models not reported here. Besides the five sets above, one further 'control' set of twenty unread documents was used as a baseline

Taking first the issue of cover data, the initial class test gave a strongly significant result, indicating that cover inconsistency appeared at different degrees in the five sets of documents. Retesting at the individual forms of inconsistency produced differing results. The presence of a frontispiece rather than the original cover in the ebook view did not prove significant (*p=0.67*), while different covers at the catalogue and ebook splash view narrowly did (*p=0.021, χ^2=13.26, df=5*), and both minor and major inconsistencies between the splash page and ebook view demonstrated clear effects on reader behaviour (*p<0.015* in both cases, χ^2=15.836, 28.08 respectively). The initial absence of a cover image at the catalogue level proved to have no clear impact on user behaviour (*p=0.997*). Problem rates are seen in Table 1, above, with an error rate over 1 showing many books in the very short group demonstrated more than one problem. In all cases that held statistical significance, inconsistency was most strongly associated with a very short or short assessment pattern, suggesting that readers may well be engaging with these items out of confusion, rather than because they had necessarily been deemed to be useful.

	Corrupt ToC	Flattened hierarchy	Toc in all capitals	Partial ToC only	Total ToC errors
Not loaned	0.00	0.08	0.14	0.12	0.34
Very short reading	0.03	0.23	0.39	0.23	0.87
Short reading	0.05	0.19	0.28	0.16	0.70
Medium reading	0.05	0.10	0.18	0.05	0.38
Long reading	0.05	0.05	0.28	0.00	0.48

Table 2: Table of contents error rates

The second class of problems surrounded the table-of-contents that accompanied each ebook. Error rates for different types of problem in this class are shown in Table 2 above. This set of features proved significant overall (*p=0.018, χ2=14.99, df=5*), but appears to be comprised of a range of weak effects, with individually non-significant values (*p* ranging from 0.09 to 0.32), though the direction of these effects is consistent. Again, inconsistency or errors are associated with short and very short reading patterns.

Within-Document Analysis

As well as considering the behaviours shown in a group of documents, we can also analyse the different behaviours exhibited on each specific document. The results of this evaluation revealed a different picture.

Starting from our global set of 900 unique books, including non-loaned items, the overall pattern of the distribution follows Zipf's (or Bradford's) law. This is reassuring, as it is the common distribution for patron loans and inspections within libraries (see for example (Christianson 2005)).

One critical point is that of the 900 unique books inspected, 136 were loaned, but never inspected without being loaned. All other books were inspected without being loaned on at least one occasion, with the tailmost item being inspected 27 times, but only loaned on one of those occasions.

Having established the general distribution, we then inspected the occurrences of the books that were included in the sets already reported in this analysis. There were 172 unique books that were represented in one or more of the four sets in our initial analysis. The pattern of this distribution was not, as might have been expected, Zipfian. We noted earlier the number of loans in each set, but the number of unique books was smaller, as see in Table 3.

	Very Short	Short Reading	Medium Reading	Long Reading	Total
Loans	33	49	64	68	214
Books	32	45	59	64	172

Table 3: Unique book and loan counts

This distribution is reproduced when represented by the mean viewing time or page counts for each title – i.e. it is not an effect of our initial selection of sets. The distribution was also not linear – rather there was a bias towards longer reading times in general.

As can be seen from Table 3, the number of books that appears in more than one set is relatively small – only 26 titles. Some caution therefore has to be applied in interpreting a rather sparse set from our original data, but this set is distributed in proportion to the different sets (χ^2 produces p=0.682, clearly rejecting any notable pattern).

Returning to the book perspective of the data, we can examine these recurring items, in terms of cover inconsistencies and table-of-contents problems. Twenty of the books were included in our main samples, and from this set the gross rate of cover errors was 0.6 – i.e. nearly identical to the 'medium reading' set. Likewise the rate of table-of-content errors was 0.35 – similar to the medium reading set again and, indeed, lower than the long reading group.

Were this set made predominantly from members of the long and medium documents, this pattern would be unremarkable. However, the number of books that occurred in each of the four reading sets was 8 (very short), 14 (short), 8 (medium), and 11 (long). The combinations of set membership also do not show any bias.

We noted earlier an association between higher error rates and shorter reading times. The data from multiply read books might therefore suggest our initial conclusion may be in error. However, there is a direct and simple common trait shared by the books that only received short

readings in this set: they lacked cover images in the main library catalogue.

The shorter reading sets do have a higher proportion of books with an icon, rather than cover, in the main library catalogue, though this is not statistically significant. However, the analysis of multiple readings does suggest that, in fact, some short readings are caused by incomplete catalogue information for books with low rates of production errors. The corollary of that possibility would, in turn, be to increase the proportion of defects in other books that received short readings.

Therefore, this detailed analysis by book indicates that further work may strengthen our primary findings regarding the impact of cover and table-of-content defects on shortened reading times. We would repeat our caution that this is a relatively small set, but given that it otherwise appears to be representative, it does appear that this tentative conclusion is worthy of consideration.

DISCUSSION

There are three areas of research relevant to our findings: document selection in digital environments, book selection in physical environments, and document triage practices.

Digital document selection practices

We can draw on studies from the field of digital libraries that identify the impact of presentation factors on the acceptability and acceptance of digital information resources.

Theng et al (Theng et al. 1999) reported that inconsistency in the presentation of classification structures in digital libraries and inconsistent document presentation negatively impacted on usability assessments of digital libraries. Digital libraries contain many digital documents, and their classification schemes are analogous to tables-of-contents within books. Similarly, the relationship between individual search results in a digital library and the documents they link to reflects the same relationship and access process for ebooks in an online library catalogue. Our discoveries reflect the same negative responses to inconsistency seen in Theng's work with books rather than an available selection of digital documents.

Later work by Blandford et al (Blandford et al. 2001) on the use of multiple digital libraries underlined the same point, indicating that acceptance of individual sources was influenced by the same factors.

This suggests that, in fact, users are sensitive to the consistent presentation and representation of data in electronic settings in general. Similar problems arise when either digital library systems or the (digital) books that they contain suffer incoherent or contradictory appearance for either content or access methods.

We can also derive understanding from the use of traditional libraries.

Physical book selection

Reutzel and Gali (Reutzel and Gali 1998) and Moore (Moore 1995) studied children selecting books in physical environments. Both studies demonstrated that shelf location influenced book selection behaviour, usually by convenience (books at eye level and the books children came across first were more likely to be selected than other books). Borgman's work with children and electronic shelf presentations suggests that this behaviour is also present in digital environments (Borgman, Hirsh et al. 1995). The analogue for this in our study is location in search results: while it is a near certainty that position in search results will affect the number of readers who navigate from the catalogue to an ebook (see for example search result data in (Jansen and Spink 2006)), it does not appear to affect interaction thereafter.

Stelmaszewska (Stelmaszewska and Blandford 2004), and McKay (McKay 2011) found that shelf location also played a part in adult information seeking in physical environments, but rather than convenience it was affected by serendipity: readers used the co-location of similar books on library bookshelves to locate other items of interest. This co-location strategy is not available to those seeking ebooks, a likely detriment to their being able to determine the usefulness of a book prior to entering the book interface as they cannot examine nearby books to determine whether they are on the right track.

Hinze (Hinze, McKay et al. 2012) and Ooi (Ooi 2008) also noticed a shelf-location effect in adults, but this effect was slightly different than the effects outlined above: They noticed readers were inclined to select books that they could see the covers of, for example display books. Reutzel (Reutzel and Gali 1998), Hinze (Hinze, McKay et al. 2012), and Stelmaszewska (Stelmaszewska and Blandford 2004) also noticed that cover detail affected book selection practices in libraries, and Buchanan found the same thing in bookshops (Buchanan and McKay 2011). We hoped with this study to identify some of the cover elements that affect book selection, however our study does not provide evidence for effects based on cover content (which is not to say these effects necessarily do not exist, just that our study did not identify them). Our study did reinforce the importance of cover, however: cover inconsistencies resulted in longer investigation times by readers without further reading.

Another facet of books identified to be important in book selection in physical environments is the Table of Contents. One study showed this to be the primary driving factor in choosing between two similar books (Stieve and Schoen 2006); Hinze (Hinze, McKay et al. 2012) et al frequently witnessed readers consulting the ToC when making decisions about books. As with cover images our study does not identify what elements within the ToC affect decision making, however inconsistency or errors in the ToC again leads to extended investigation of books without significant reading.

Our work, therefore, reinforces the importance of two elements (cover and table of contents) in the digital space previously identified as important in the physical space, though not in the ways in which we might expect. Rather than identifying information that predicts use, we have identified areas where inconsistencies predict behaviour

that appears to represent reader confusion, an impression reinforced by the literature on document triage.

Document triage

In previous work (Loizides et al. 2009) it was reported that when users were assessing documents for their relevance to a search task, shorter reading times were associated with clear relevance decisions. When a document's relevance was ambiguous or uncertain, more of the document would be examined over a longer time. This outcome mirrors our own finding - that documents with clearer and more consistent indicators of their content would either be examined briefly, or read over an extensive period of time. In contrast, documents that are presented inconsistently are read for short periods of time, but seldom just rejected rapidly, and rarely progress to being read extensively. This suggests that our records repeat the pattern of Loizides' initial assessment, with the addition of longer focussed reading of selected documents.

In a separate study (Loizides 2007), Loizides noted the importance of the first page of electronic documents in document triage, particularly in comparison to paper. Our study reinforces this, as the elements we identified that lead to confused behaviour (cover image inconsistencies and ToC problems) are visible from the first page of an ebook. Reducing the errors in these elements would make it easier for readers to determine which ebooks are useful without having to engage further.

Bae et al (Bae et al. 2006) noted that early intense reading was heavily associated with document usefulness in online triage situations, though reading patterns during triage were not consistent between users. In all cases, however, reading tailed off toward the end of document interactions. This reinforces our findings: in our sample brief initial reading was associated with confusion, rather than actual document usefulness.

In all three studies mentioned here the target documents were online journal articles; it is interesting to note that our findings show similar behaviour in ebooks.

CONCLUSIONS

In this paper we present a close examination of the interface elements readers see during the ebook triage process, and their effect on the length of reader interactions with ebooks.

We anticipated identifying certain types of cover, blurb and table of contents that would encourage readers to create loans and interact with books. While we did find that these elements affected book interaction behaviour, it was not in the way that we expected. Rather than certain types of cover or types of information in table of contents encouraging longer interaction, we discovered that inconsistencies or errors in this information encouraged a certain specific interaction pattern: loan, interaction and abandonment of an information resource.

This pattern holds true whether we look at loans, or interactions with specific documents: those documents with inconsistencies or errors are more likely to be seen only in the group of books that users loaned and interacted with briefly before abandoning them.

Given that a loan is the condition under which libraries pay for access to books this pattern should be of considerable concern to libraries, and data improvement should be a priority to maximise beneficial use of resources.

Our study does open some avenues for future work: it is not clear from the logs how many people reject ebooks in the catalogue view, or at the ebook splash page. Identifying reasons for rejection or acceptance at this point could provide more insight into the value of cover, blurb, title and other bibliographic data to ebook selection practices. The transaction logs we have do provide us with information about behaviour by individual users that was not taken into account during this study; in some instances users appear to have examined a series of ebooks. Studying behaviour at a user level could also provide further insight into the user interface elements and content of books with which readers engage most heavily.

ACKNOWLEDGMENTS

We would like to thank Tony Davies at Swinburne University of Technology for his help identifying and understanding EBL data, and Alison Morin at EBL for her continued support of our ebook research.

REFERENCES

Agosto, D. E. Bounded rationality and satisficing in young people's web based decision making. Journal of the American Society for Information Science and Technology 53, 1 (2002) 16-27.

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.

Ballard, T. and Blaine, A. User search-limiting behavior in online catalogs: Comparing classic catalog use to search behavior in next-generation catalogs. New Library World 112, 5 (2011) 261-73.

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) 518-25.

Blandford, A., Stelmaszewska, H. and Bryan-Kinns, N. Use of multiple digital libraries: A case study. Proc. Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries, ACM (2001), 179-88.

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. Journal of the American Society for Information Science 46, 9 (1995) 663-84.

Buchanan, G. and McKay, D. In the bookshop: Examining popular search strategies. Proc. JCDL 2011, ACM (2011),

Casey, S. M. Set phasers on stun: And other true tales of design, technology, and human error Santa Barbara, CA, USA, Aegean, (1998).

Christianson, M. Patterns of use of electronic books. Library Collections, Acquisitions, and Technical Services 29, 4 (2005) 351-63.

Connaway, L. S., Dickey, T. J. and Radford, M. L. "If it is too inconvenient i'm not going after it:" Convenience as a critical factor in information-seeking behaviors. Library & Information Science Research 33, 3 (2011) 179-90.

Cooper, M. D. Usage patterns of a web-based library catalog. Journal of the American Society for Information Science and Technology 52, 2 (2001) 137-48.

Cousins, S. A. In their own words: An examination of catalogue users' subject queries. J. Inf. Sci. 18, 5 (1992) 329-41.

Crabtree, A., Twidale, M. B., O'Brien, J. and Nichols, D. M. Talking in the library: Implications for the design of digital libraries. Proc. 2nd International ACM Conference on Digital Libraries, ACM Press (1997), 221-8.

Cummings, I. (2011). Kansas libraries lead way in e-book access. Kansan.com. Kansas, USA.

Dix, A., Finlay, J. and Abowd, G. Human-computer interaction, Pearson/Prentice Hall, (2004).

Glaser, B. G. and Strauss, A. L. The discovery of grounded theory: Strategies for qualitative research. Hawthorne, NY, Aldine de Gruyter, (1967).

Hamblen (2011). Amazon: E-books now outsell print books. Computerworld. 2011.

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.

Hinze, A., McKay, D., Vanderschantz, N., Timpany, C. and Cunningham, S. J. Book selection behavior in the physical library: Implications for ebook collections. Proc. JCDL '12, ACM (2012), 305-14.

Holt, K. (2010). E-book sales statistics from bisg survey. Publishing Perspectives. 2011.

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) 248-63.

Jones, S., Cunningham, S. J. and McNab, R. An analysis of usage of a digital library. Proc. ECDL'98, Springer (1998), 261-77.

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, Springer, (2011).

Liesaputra, V. and Witten, I. 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) 256-62.

Loizides, F. Investigating document triage on paper and electronic media. Proc. ECDL 2007, Springer Berlin / Heidelberg (2007), 416-27.

Loizides, F. and Buchanan, G. An empirical study of user navigation during document triage. Proc. ECDL 2009, Springer Berlin / Heidelberg (2009), 138-49.

Mahoui, M. and Cunningham, S. J. A comparative transaction log analysis of two computing collections. Proc. Research and Advanced Technology for Digital Libraries: 4th European Conference, ECDL 2000, Springer (2000), 418-23.

Mahoui, M. and Cunningham, S. J. Search behavior in a research oriented digital library. Proc. Research and Advanced Technology for Digital Libraries: 5th European Conference, ECDL 2001, Springer (2001), 13-24.

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), 225-34.

McKay, D. Gotta keep 'em separated: Why the single search box may not be right for libraries. Proc. CHINZ '11, ACM (2011), 109-12.

McKay, D. A jump to the left (and then a step to the right): reading practices within academic ebooks. Proc. OZCHI '11 ACM (2011) 202-210.

McKay, D., Hinze, A., Heese, R., Vanderschantz, N., Timpany, C. and Cunningham, S. J. An Exploration of ebook Selection Behavior in Academic Library Collections. Proc. TPDL '12, Springer (2012), 13-24.

Moore, P. Information problem solving: A wider view of library skills. Contemporary Educational Psychology 20, 1 (1995) 1-31.

Nordlie, R. "User revealment"— a comparison of initial queries and ensuing question development in online searching and in human reference interactions. Proc. 22nd Annual ACM Conference on Research and Development in Information Retrieval, ACM Press (1999), 11-8.

Ooi, K. How adult fiction readers select fiction books in public libraries: A study of information seeking in context. Wellington, New Zealand, Master of Library and Information Studies Thesis completed for Victoria University of Wellington, (2008).

Pearson, J., Buchanan, G. and Thimbleby, H. HCI design principles for ereaders. Proc. Booksonline '10, ACM (2010),

Ponsford, B. C. and vanDuinkerken, W. User expectations in the time of google: Usability testing of federated searching. Internet Reference Services Quarterly 12, 1 (2007) 159 - 78.

Reutzel, D. R. and Gali, K. The art of children's book selection: A labyrinth unexplored. Reading Psychology 19, 1 (1998) 3-50.

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.

Salvucci, D. D. Predicting the effects of in-car interfaces on driver behavior using a cognitive architecture. Proc. Proceedings of the SIGCHI conference on Human factors in computing systems, ACM (2001), 120-7.

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.

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),

Theng, Y., Duncker, E., Mohd-Nasir, N., Buchanan, G. and Thimbleby, H. Design guidelines and user-centred digital libraries. Proc. ECDL '99, Springer (1999), 852-3.