McKay, D. & Buchanan, G. (2015). You Can Check It Out But It Will Never Leave: Characterising Ebook Borrowing Patterns. Paper presented at the ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR), 13-03-2016 - 17-03-2016, North Carolina, USA.

You Can Check It Out But It Will Never Leave: Characterising Ebook Borrowing Patterns

Dana McKay
University of Melbourne
Parkville 3010
Australia
dmckay1@student.unimelb.edu.au

George Buchanan
Centre for HCI Design, City University London
Northampton Square
London, EC1V 0HB
+44 20 7040 8469
george.buchanan.1@city.ac.uk

ABSTRACT
What does it mean for a reader to borrow an ebook? Ebook technology means that borrowing can take different forms, for example printing and reading. We do not know, though, which of these options readers actually use. Ebook technology generates logs that allow us to understand ebook borrowing patterns over time, both by individual readers and in aggregate. Despite the ready availability of ebook logs, this area remains under-researched. In this paper we present an exploratory log analysis of ebook borrowing, comparing printing and reading, discovery patterns, single- and multiple-book sessions and identifying specific borrowing patterns.

Categories and Subject Descriptors
H.5.m Information interfaces and presentation (e.g. HCI) miscellaneous

General Terms
Measurement, Documentation, Design, Human Factors

Keywords
Ebooks, log analysis, reading, information behaviour.

1. INTRODUCTION
When a reader borrows a print book from a library, they check it out, and—we assume—read it before returning it. The book is a physical thing, and while it is checked out to one reader cannot be borrowed by another. Similarly individual readers (usually) leave little trace of what they did with borrowed books, though cumulative wear will appear over time. Books are borrowed from shelves, where—in academic libraries at least—they are shelved near books on similar topics. Within library opening hours, readers may sample books in the library for any length of time, even completing a reading. Finally, with the exception of copying for later, books can only be read in one medium—the printed pages they contain.

As Marshall rightly notes, ebooks represent a different proposition [35]. They can have numerous simultaneous users, and never show any signs of wear. Ebook providers typically require readers to take out a loan within a short time of beginning to read, though this process is largely transparent to readers. Ebooks are not shelved at all, and generally only accessible using search: ebook browsing capabilities, especially those for academic ebooks, are relatively poor. Most importantly for this paper, though, ebook providers can (and often do) record all ebook usage, even down to the page level. This data allows us to study ebook use in a way that would be so obtrusive as to be 'creepy' [35] to do for print books.

Data on ebook use has been leveraged at a basic level, both to allow patron driven acquisition [13] and to compare raw loans between print and ebooks [7; 8]. It has also been used less widely to study how readers triage ebooks [39; 40], and once to study reading behaviour [37]. The capacity for logging, though, has not (to our knowledge) been used to study what it means to borrow an ebook given the technological differences between ebooks and print books outlined above. There is also a dearth of literature on ebook borrowing patterns (outside of a small number of comparisons with print borrowing [8; 29; 42]). In this paper we aim to understand how reading and printing interact in ebook borrowing; how individual readers borrow over time, and what the differences are between sessions where readers borrow a single ebook and sessions where they borrow more than one. By looking at these facets of ebook use we hope to characterise what it means to borrow an ebook.

This paper is organised as follows: Section 2 addresses the background literature, Section 3 outlines our research method. Section 4 compares printing and reading in ebook borrowing, and Section 5 describes the differences between sessions where readers borrow a single ebook, and those where they borrow many. Section 6 discusses these results in view of the literature in the field; finally we draw conclusions and provide avenues for future work in Section 7.

2. BACKGROUND LITERATURE
This paper builds on both the literature on book borrowing (Section 2.1), and the literature on ebook usage (Section 2.2). It also builds on the literature on reading—both ebooks and print books (Sections 2.3 and 2.4 respectively).

2.1 Book Borrowing
There is a long but sparse history of studying print borrowing in libraries. The most common traditional method has been the materials availability survey, where library users are asked whether they found what they were looking for [22]. The advent of online cataloguing allowed more detailed analysis of borrowing, though apart from examining users' search behaviour

CHIIR '16, March 13 - 17, 2016, Carrboro, NC, USA

[11; 26; 61] this was poorly leveraged until recently. One early study based on borrowing logs exists; it found that relative shelf location (and thus browsing) was an strong predictor of borrowing for a few small subject areas, at least [30]. This assertion is supported by a contemporaneous study of the impact of a new online catalogue on reader behaviour, which found that over 50% of users who found one book they had identified in the catalogue on the shelves also borrowed at least one further book [12].

More recent studies of borrowing behaviour have tended to use an observational approach [16; 17; 31; 56] or interviews [48; 53]. These studies have typically investigated how readers find books rather than aiming to understand broader borrowing patterns.

Finally, we have recently used transaction log analysis to demonstrate borrowing patterns more widely. An initial study found that proximal books were more likely than distant books to be borrowed together [41]; this was followed up by a study demonstrating that browsing appears to be a better predictor of borrowing than search, that day of the week affects borrowing rates, and that books are more likely to be borrowed together within classification boundaries than across them [38].

The one finding common to studies of print borrowing is the prevalence and importance of shelf location in borrowing patterns: readers value the opportunity to browse [5; 17; 31; 51], and log analysis demonstrates they leverage it willingly [38; 41; 42]. This is not to say that the shelves are perfect—readers are well aware that books can be checked out or misshelved [5]. Even so, given the dominance of shelves as an influencer in print borrowing, what does this mean for ebooks, where—dependent on one's viewpoint—either there are no shelves, or they are infinitely re-arrangeable? We aim to address this question in this paper.

2.2 Ebook usage

The equivalent studies to the browsing studies above are ebook usage studies. These studies typically have been used to establish that ebooks are used at all, rather than how they are used. In this vein, early studies of ebook usage typically compared ebook circulation with print circulation [8; 29]. These circulation comparisons found that ebooks were used more than equivalent print books in some disciplines—usually tech and science related—but not others. A similar study that looked solely at the use of ebooks found similar discipline effects, and that ebooks were used in a power-law distribution—some were very popular, but popularity tails off exponentially [7].

These studies were followed up by a spate of reader surveys to find out what proportion of users was using ebooks, and—where possible—whether these groups had any identifying characteristics. These studies found a number of things—undergraduates were more likely to prefer ebooks than other users, as were men and those in technical fields [27; 52]. Similar studies asked users what advantages and disadvantages they saw in ebooks [15; 52; 54]; results were fairly consistent. Users appreciated the ability to search within ebook content anytime anywhere access and but were irritated by DRM and usage restrictions and poor annotation capacity.

Finally our own very recent study of ebook and print book use compared borrowing patterns [42]. This study found that ebooks were more likely to be borrowed in single-book transactions than print and that groups were more loosely topic clustered. The total number of unique books borrowed by each user was marginally higher for ebooks than print, suggesting convenience is a factor in book use.

None of these studies address what actually happens during an ebook loan, however, nor whether there are patterns of behaviour with repeated loans. We aim to address this gap in this paper.

2.3 Reading Ebooks

Studies of how people read ebooks can be roughly divided into three groups: usability studies, experimental studies, diary or interview studies of some particular aspect of ebook reading, and log studies.

Usability studies of ebook reading have compared print and screen reading, or focused on ereader technology. Comparisons of print and screen reading initially suggested that screen reading was more difficult than print [10], though this difference has been ameliorated by the introduction of e-ink technology [35]. Usability studies of ereaders have identified a number of issues, including poor navigation and annotation capabilities [47; 50]. Ereaders are still more usable than attempting to read from the screen of a larger device, however [35].

Experimental studies have examined a number of aspects of ebook reading. Berg [4] and Malama [32] both studied navigation within ebooks, and found that readers used table of contents and index frequently, but that they struggled with ebook navigation. Liesaputra et al noted that a realistic book presentation somewhat ameliorates these difficulties [28]. Thayer et al gave readers the choice of print or e-textbooks, and found that print was generally preferred, but that ebooks were more frequently used on the move [59].Takano and colleagues have looked at a range of aspects of ereading in comparison with print (for example [55; 57; 58]), notably document handling and navigation. These studies have consistently found print to be more usable, particularly for navigation. Finally, Mangen compared recall in readers of a print novel and its ebook equivalent, and found almost no difference [33]. The one exception was event order in the story: print readers recalled this better than ebook readers. These studies focus on individual aspects of ereading, however, and do not allow for the long term examination that ebook logging technology affords.

Two studies have used logs to examine ebook selection, though not deep reading [39; 40]. These studies found that triage behaviour looks similar in ebooks to print books, and that cover and table-of-contents are important in book selection. Finally, two studies have used logs to address reading specifically. The first examined navigation during reading, and found significant flicking backwards and forwards and some use of the electronic table of contents [37]. The other dates from 2009 [9], and reports that the typical ebook reading session covers only a handful of pages (usually fewer than 7) in a few minutes.

No study has harnessed the power of ebook logging to examine not just reading, but printing, nor has any study attempted to use logs to characterise user behaviour over time. This paper attempts to address some of these gaps.

2.4 Print Reading

Marshall rightly notes that it is difficult and to study long form print reading [35], however the use of videotape and diary studies have allowed some insight into this area.

Video recordings of reading [36] show that readers move around a lot while reading print, that they flick backwards and forwards within a text regularly, and that they use tables of contents—the latter two have also been seen in ebooks, as noted above.

A diary study of work-related reading notes that there are a number of types of reading, including deep reading, scanning, and

fact checking [1]. This work also notes that annotation and navigation are core parts of these activities.

These studies are fairly representative of the studies of print reading, and identify the major problem with studying it: it is simply not possible to do unobtrusively in the way that ebook logs allow. For the first time in history ebook logs allow us to watch readers at work; this paper leverages that capability to further our understanding of academic reading patterns.

3. METHOD

In this section we first describe our dataset, then define the meaning of an ebook loan, and then outline our approach to analysing this data.

3.1 The Dataset

This paper is based on two sets of ebook log data provided by the library at Swinburne University of Technology, a small, research-active Australian university. The datasets are from 2013 and 2015, each from April to July (inclusive); this is core term time in Australia.

The collection's technology and access policy has been stable over two years; hence changes will primarily reflect changes in behaviour, rather than alterations in response to, for example, changes in search interfaces or the reading experience.

The system being used is EBL, an established ebook platform used in many academic libraries. The interface for this platform is shown in Figure 1. The configuration of EBL at Swinburne does not allow for PDF downloads of ebooks. Readers must instead view ebooks online, print them or copy content. The system logs user behaviour anonymously, using a patron identifier that is unique to each user, but cannot be de-anonymised. We can thus track the behaviour of individual readers over time.

Each log entry reflects a single ebook loan. Loans only occur after one of a number of conditions is met. Most commonly, this is because the reader views the book for longer than a set period (ten minutes for books pre-purchased by the library, or five minutes for other books). Other loan conditions include reading or more than 20% of the pages of a book, copying (via cut-and-paste) part of the content of a page, or printing one or more pages. Each log entry records the book, patron identifier and a range of information, including the date and time of the loan, the duration of user activity, the quantity and page numbers of any pages viewed, printed or copied, and other bibliographic information.

Loans are not automatically generated; readers must confirm (by clicking 'yes' in a click through box) that they wish to borrow a book. This confirmation is lightweight—it requires only a single click, and does not have any financial implications for readers.

We cleaned the data prior to analysis; removing 1281 loans that recorded no usage activity, and over 70 loans that were exact duplicates of other loans.

3.2 Our analysis

Our overall methods draw from previous log-based analyses of library and information behaviour, for example [7; 21; 26; 41; 45]. To avoid over-testing, and allow for the evaluation of interaction between factors, we used log-linear analysis—a non-parametric test for population frequencies.

Our first aim was to understand the contrasting role of printing versus online viewing, and the degree to which these were co-occurring actions. A high reliance on print would indicate that ebooks are primarily used as a fast delivery method for paper,

Figure 1. The ebook reading interface. Note the interactive table of contents at left.

whereas a high level of online viewing would be more suggestive of a 'born digital' ebook culture. Different researchers have argued passionately for opposing viewpoints [20; 43], but we lack data to support either view. We similarly do not know if printing and viewing are of the same material, or of distinct content.

Examining individual page sequences of views can further reveal the degree to which reading is either sequential or non-linear. Again, our current research data is sparse. Though Marshall observes that strictly linear reading is rare even in print [36], non-linear reading is both more common and more random online [37].

As with any novel technology, practise will vary in early adoption. With access to data from two separate years, we could identify changes in behaviour across time. Determining which factors are stable and which are in flux gives us clearer insight into user behaviour.

To provide a foundation for our analysis, we established a global view of the data, counting the number of loans that included every possible combination of printing, viewing and copying to identify the relative frequency of these actions. We grouped loans that occurred within 30 minutes of each other and had a single reader to form co-borrowing sessions; in line with our earlier work [42]. All other loans were singleton loan sessions. For both co-borrowing and singleton loans the frequency of printing and viewing were established. Where a loan included both, we tested for any overlap in the pages used, to determine if these were interconnected or distinct activities.

One property of digital books is the potential for direct copying. As the public data on digital copying is minimal, we tallied the volume of both the number of loans that included copying, and the number of pages copied. The rates of copying in co-borrowing and singleton sessions were counted separately. We also counted the proportion of copying that was accompanied by printing and viewing.

We then embarked on three separate analyses, comparing co-borrowing versus single-book sessions; second and subsequent readings of the same book by individual users; and book-specific patterns across all readers. For each ebook, we measured:

- The total duration of a loan
- The total number of pages viewed, their sequence both in time and in page order, and the number of times each was viewed.
- The total pages printed, and again sequence and frequency.
- The number of pages copied

First, we compared the co-borrowing sessions against single-book sessions to identify any differences in the four measures above.

Previous research on reading has reported that reading of pages is often non-linear, with considerable skipping of pages. For both co-borrowing and singleton loans, we established the span of

pages both viewed and printed. These measures were then compared by loan type and year.

Second, we examined renewal (repeat) readings of individual ebooks by single users, again using the four measures above. This would show if renewal readings differed with each renewal. We gathered the profile for the co-borrowings in each sequence, to determine if the earlier loans would vary from the later ones in terms of time, printing and viewing preferences, etc. Again, 20 individual books were analysed by hand.

Third, we aggregated all the sessions for each ebook, and identified which of its pages that were repeatedly used by different readers for printing, viewing and copying. This enabled us to see if different readers used the same ebook in similar ways, and determine commonly accessed material. We sampled the 100 most loaned ebooks and collated the number of times each page was printed or viewed. This was further consolidated into contiguous groups of pages that were used the same number of times. A subset of 20 was manually inspected to identify the content of the pages in terms of the book structure (chapters, front matter) and page content (tables, diagrams etc.)

To conclude the analysis, we finally turned to assess the impact of topic on co-borrowing. We created a further sample set that excluded any book that was loaned fewer than ten times (to minimise data noise). This set allowed us to investigate the impact of (a lack of) physical shelves on borrowing patterns. Our recent work [38; 41; 42] strongly suggests that browsing on physical shelves increases co-borrowing (borrowing more than one book at a time) where books are shelved close together. Topic is a confounding factor in shelf proximity; academic library shelves are organised according to topic. Our recent comparison of ebook and print borrowing has shown that ebook loan clustering around topic occurs, but that it is looser than print clustering [42]. To further understand this phenomenon, we created a virtual shelf of the ebooks available in our collection sorted by Dewey decimal number, then by title (sadly the author metadata needed to replicate library practice was not available). We examined ebook co-borrowings according to their nominal shelf distance to assess the impact of topic (clearly relevant in ebook borrowing) versus browsing, which does not exist for this collection.

4. RESULTS

From the 40708 loans in the two years, 36064 (89%) included viewing of pages, 10170 (25%) included printing, and 4034 (10%) included the copying of content from one or more pages. A small residue of 281 loans (0.7%) recorded only the copying of one or more pages, but no viewing of pages. 1281 loans recorded no activity whatsoever, and these were omitted from these totals.

Our first interest was the relationship between the volume of pages users viewed and printed. For the overall set, there was no significant correlation between the number of pages printed and viewed in a single loan ($r=0.06$). If we examine only those loans where both printing and viewing took place, there is a moderately significant correlation between the number of printed and viewed pages ($r=0.39$, $df=4541$, $p<0.001$).

There was no significant correlation between the number of pages where content was copied in a loan and the number of pages accessed in any other way.

4.1 Co-borrowing and single- ebook sessions

We separated loans into two sets: singleton loans, where readers accessed only a single book, and co-borrowing sessions, where the readers accessed two or more ebooks each within 30 minutes of finishing the last.

Table 1. Patterns of loans: single and co-borrowing sessions

Loans	2013	2015	Total
Co-borrowing sessions	856	2161	3017
Co-borrowing ebook loans	1864	4898	6812
Single loans	11623	22274	33897
Total loans	13487	27222	40709
Total sessions	12479	24435	36914
Unique users	4945	6752	11697

This basic summary revealed both that the number of both sessions and loans nearly doubled across the two years, and the number of unique users rose by over 35%. While the number of ebooks loaned per session was relatively stable, the number of sessions and loans per user increased.

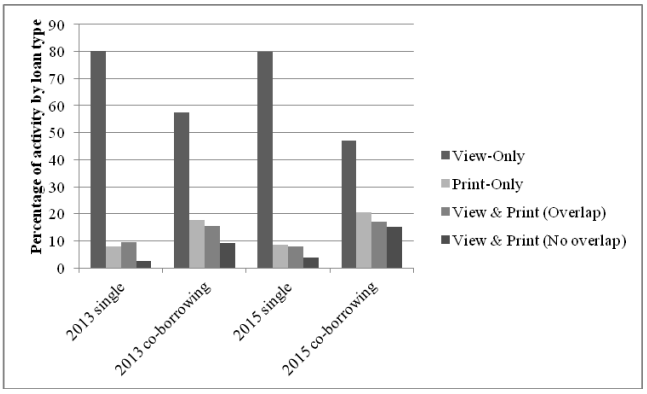

Figure 2: Types of use activity according to loan type

We then examined whether singleton and co-borrowing sessions were associated with different types of user activity. There was a marked difference in the balance of viewing and printing activity (see Figure 2). It is clear from our data that co-borrowing involves a much higher degree of printing than singleton borrowing, behaviour that looks somewhat like checking a number of books out of a library to read later. There is also a slight increase in printing in 2015 over 2013. This data shows that despite the cost of printing to users (where viewing is free), there are clearly times where the goal of ebook activity is to locate and print content from a number of books.

Analysing the data using a log-linear test across the three factors (loan activity, year and session type), produced $p<0.0001$ ($G^2=2853.61$, $df=13$). The strongest interaction was between session type and loan activity, ($G^2 = 2577.86$, $df=4$), but all interactions were significant ($G^2=157.38$, $df=4$; $G^2=126.1$, $df=1$; $p<0.0001$).

4.2 Copying Content

The collection permits copying of small portions of a book, creating a loan when readers do so. Between 2013 and 2015, while the total volume of copying was similar, the proportion of sessions that included copying fell markedly (see Table 2).

The likelihood of copying was not affected by the temporal length of loans, but the number of pages copied increased with longer reading. Loans under 5 minutes that included any copying had an average of 1.6 pages copied; this rose to 2.4 pages by 10 minutes, and quickly plateaued at around 3 pages.

Table 2. Copying rates and activity

	2013		2015	
Loan Type	**Single**	**Co-Borrow**	**Single**	**Co-Borrow**
Included Copy	33.9%	23.4%	13.0%	18.1%
Copy only	103	10	143	25
Copy & print	33	5	53	21
Copy & view	1432	185	1730	345

4.3 Session activity and length

The time readers spent on each loan decreased markedly between 2013 and 2015. The average time per book in 2013 was 29m36s for singleton loans, and 19m20s for co-borrowed books. In 2015 this fell to 17m8s and 8m40s respectively. While it may seem counterintuitive that session time shortens with co-borrowing, one explanation for this is the increase in printing with co-borrowing seen in Section 4.1; printing takes less time than reading.

Figure 3 shows the average number of pages accessed and page span (distance from first to last page accessed) for each loan type.

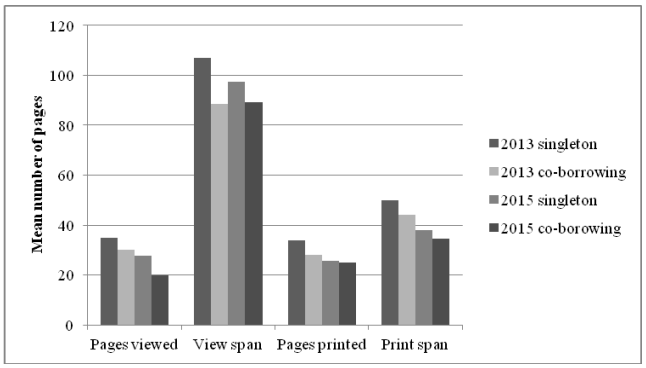

Figure 3: Number of pages viewed and printed by loan type

The most obvious feature of this graph is that the span of pages accessed is much greater than the number of pages accessed—readers are skipping past a great deal of content, in common with earlier work in the area [37]. For printing, the number of skipped pages is smaller than for viewing, this discrepancy is likely to be a result of the viewing interface presenting single-click options for skipping multiple pages.

The next thing to note is that for co-borrowings the span viewed and printed is largely the same in 2015 as 2013, but the number actually accessed dropped substantially: users are examining a smaller number of pages within a similar span.

4.4 Viewed pages per loan

Over 30% of viewing loans started on the first page, however a similar number started beyond the 33rd page. This distribution is, as with so many information behaviours, Zipfian. This distribution also applies to the span of pages covered by each reading session: 33% are shorter than 18 pages, and the same proportion longer than 101. The later in the ebook viewing began, the fewer pages were read, a correlation which is marginally significant (r^2=-0.2).

Viewing was primarily linear, but perfectly continuous reading was rare, as has been observed in print and previous studies of ebooks [36; 37]. While there were many jumps of 4-5 pages (mean 4.53 pages for all loans), the mean longest jump per book was over 50 pages for both years, a result also seen in 2011 [37]. These big leaps explain most of the difference between the span of pages and the smaller number of pages read seen in Figure 3.

We also analysed different activity factors—total page counts, and time per page—by time, initially separating sessions by bands of five minutes.

Loans under 5 minutes (close to 50% of all loans in both years) were less likely to be linear than longer ones; the mean number of consecutive pages viewed was 2.3 in 2013 and 2.55 in 2015.Sessions that exceeded 10 minutes typically included sequences of 8 pages (mean 7.83 pages) and when reading lasted over an hour, runs of 15 pages or more were commonplace.

The average length of time spent on each page echoes this difference. For loans under 5 minutes the mean was a bare 6.4s per page; for loans over 15 minutes, this was 28.75s and beyond an hour, the mean time per page rises to 93.14s. It is quite likely that the latter figure includes a much higher proportion of idle time than with a short viewing time.

Bearing in mind that loans generated by the library only commence after a large number of initial pages, or several minutes, short loans are perhaps surprisingly more similar to triage-like skim reading than might be expected. Conversely long loans begin to look like deep reading.

4.5 Patterns in co-borrowing

We wanted to know whether readers' behaviour changed as they moved through the books borrowed in a co-borrowing sequence. We were particularly interested in the proportion of printing and viewing over time; after all co-borrowing could just be long-form triage or there may be some other pattern. To understand the nature of co-borrowing, we examined both the overall behaviour in borrowing sequences and the behaviour in the final book in particular. Results are shown in Figure 4.

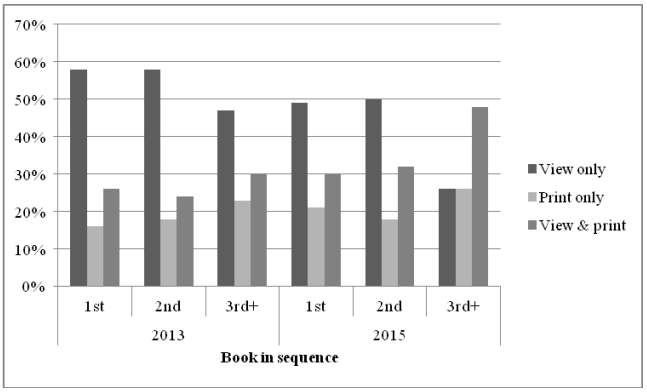

Figure 4: Actions by order of co-borrowing sequence

There is a drop in the rate of view-only loans over time, particularly by the third or greater book in a sequence. This suggests that readers viewing three or more books are either triaging books and printing material for later, or that readers run out of time to read everything they need from the screen. It is also clear that there is more printing generally in 2015 than 2013. . Testing these differences for significance using a log-linear test, produced $p<0.0001$ (G^2=354.1; df=31), confirming that both the difference between years and the fall in view-only loans across a sequence are reliable.

When examining whether readers showed access preferences over the course of a session, we found a marked result. 95% of view-only loans were followed by another view-only loan, where random distribution would have predicted around 60%. For print-only loans, 49% were followed by another print-only loan, 49% by a print-and-view loan, and only 2% by a view-only loan. For

loans with both viewing and printing, there was less than a 10% chance of a view-only loan for the next book.

Overall, co-borrowing sessions were either printing- or viewing-focused. Manual inspection of a sample of books revealed that the number of viewed pages was limited in print-focussed sessions. Viewed pages in print session mostly appeared in the front matter, with a small number (3-4 pages) being located in the printed region, usually towards the very beginning. This appears strongly to suggest that viewing can be target-acquisition task, used identify material for printing.

In contrast to session behaviour, most users both printed and viewed ebooks over the course of their usage: The behaviour focus in sessions represents user tasks for those sessions, rather than a preference by individual users for printing or viewing.

4.6 Re-reading ebooks

Given the short loan period of ebooks in this collection (24 hours), the restrictive online reading environment (ebooks are not downloadable) and the possibility that readers were using books for reference materials it seems likely that individual readers would borrow books in this collection more than once. We examined whether this behaviour occurred, and what actions readers took with individual books over a number of re-readings.

It is clear in the data that readers return to books: every page read had a 56% chance of being re-read by the same reader in later sessions. 7825 ebooks were borrowed more than once by a single user; 58.4% of all loans were repeat borrowings (see Table 3). Repeat borrowings account for more loans than co-borrowing sessions.

There was a rise in repeat borrowings in 2015: 56% of all loans being part of multiple readings of an ebook, compared to 37.5% in 2013 ($p<0.0001$, $\chi^2=2511.2$; df=1).

Table 3. Renewed ebook behaviours

Loans	2013	2015	Total
Books re-read	2005	5280	7285
Total re-readings	5170	15874	21044
Two readings	1398	3102	4500
Largest # of re-readings	23	50	n/a

With each reading, the odds of a further reading increased: 41% of 2nd readings were followed by a third, 50% of those by a fourth, and 56.9% of those, in turn, by a fifth. The actions taken during each loan influenced this progression. The likelihood of a print-only loan being the last in a sequence was 13-16% greater than a view-only loan. For example, 56.5% of view-only second loans were the last, 71.3% of print-only loans were, a difference that is significant ($G^2=806.65$; df=28, $p<0.001$ for 3rd to 6th loans). Loans with both printing and viewing fell in the middle of this range. The high rate of termination with printing suggests again that print is being used for long-term reference or record. The number of printed pages, in contrast, was highest in the first loan—28.7 pages; later loans show an average of around 22 pages printed.

Conversely, the ratio of viewing-only loans rose from under 80% to over 90% of later readings. The average number of pages viewed grew from 23.9 pages, up to a plateau of approximately 35 pages (later re-readings varying from 34.8 to 35.3 pages). The length of sessions also increased, from 15m 28s (first reading) to 25m 55s (5th reading or later). This suggests that later readings are usually in depth, rather than brief access for reference.

Readers showed preferences for interacting with each book: viewing is disproportionately followed by viewing, and printing begets further printing. When examining variations between first and second readings with respect to activity—print-only, view-only, and print-and-view—consistency of access method is statistically significant. ($p<0.0001$, $\chi^2=1381.95$, df=4).

Readers do return to previously borrowed books, and they have behavioural preferences when doing so—viewing begets viewing, and printing begets printing. Final loans, though, show more printing than other loans—possibly readers saving material for later review or reference.

4.7 Co-borrowing sessions & re-read ebooks

Having examined behaviour in co-borrowing sessions and re-read ebooks separately, we then investigated interactions between the two. To recap, c. 15% of books are read in co-borrowing sessions, whereas repeatedly read books account for just under 60% of total loans. When examining only those books that were re-read, 48.6% of all first readings occurred as part of a co-borrowing session. The prevalence of co-borrowing decreases as the number of borrowings goes up: second readings show co-borrowing at around 33%, third readings 30%. Given the overall low rate of co-borrowing, these numbers seem higher than chance would predict.

We tested this relationship with a 3-way log-linear test of year, reading (1st to 3rd) and session type (co-borrowing or single-book), producing $p<0.0001$ ($G^2=780.3$, df=7). All 2-way interactions proved significant. In 2015, co-borrowing was more frequent in the first reading, and dropped off more by the third: suggesting that initial readings are more often part of a "triage" activity, and that if read, focus turns more onto the individual text.

4.8 Ebooks as reference material

Given the prevalence of reading lists as part of the academic experience [33; 51; 59], it seemed likely that we would be able to identify books or parts of books that were accessed in similar ways (reading a specific page or chapter, for example) by groups of readers.

To investigate this likelihood we aggregated the data of all users for each ebook, looking at which pages were printed and viewed. Using a sample of the 100 most frequently borrowed ebooks, we noted the ten most popularly viewed and printed pages in each. These lists had surprisingly little overlap: only 25% of pages were found in each list, even when omitting the first 10 pages of each book from the 'viewed' lists. This disparity was not universal, some books had a strong correlation between viewed and printed material; one such example is *Lessons on the war on terror*, all printing was between pages 92 to 120, and no page outside that range was viewed more than once.

Conversely, nearly 40% of these books had no pages that were both printed and viewed. "The Magic of Mathematics" was one—the most printed span of pages consisted of the first full chapter (p. 3-32), however, the most viewed content consisted of the fore-matter and preface, and the fourth chapter; indeed only two of the pages of the much-printed chapter (printed 51 times) were viewed by any user. Similarly, only four of the 29 pages comprising the second-most printed part of the book were ever viewed. Considering that this book was borrowed 400 times it is striking that there were no views of numerous pages that were regularly printed. In books demonstrating this pattern of use, viewing is typically confined to early pages in the book, or pages that could help readers identify the material they wish to print.

Many books showed printing that repeatedly covered one or more whole chapters. "Extending Thought in Young Children: A Parent - Teacher Partnership" was a popular text in 2015, with over 150 separate loans, 39 of which included printing. While three printed

extracts started from other pages, the other thirty-six began at the start of Chapter 5. All but three were for the chapter's full 52 pages; one omitted the last (incomplete) page, one the last five pages, and one covered only the first nine pages. Similarly, "Promoting Emotional and Social Development in Schools: A Practical Guide" had 42 prints in 185 loans, with 37 starting on the same page; and "Alone Together" by Sherry Turkle had every print run begin on the first page of Chapter 10. In some cases—just over 10% of the total prints of whole chapters—two or more contiguous chapters were printed; we coded these multi-chapter prints as a special case of the whole-chapter case.

Another recurring pattern was the printing of self-contained content within a chapter, usually comprising one or two pages. In "Sociology, the basics", patrons printed 11 individual chapters, but 16 of the 29 prints were of single pages. 14 of these were of a page containing a set of guidelines, and the other two contained a list: one of websites, one a checklist. Similarly, "SPSS Survival Guide" had 25 printing loans, of which 12 printed out a three-page set of instructions, and 15 one or more complete chapters (2 printing a table and the immediately following chapter). Printing short excerpts was more commonly found in instructional textbooks, and reference material. In both types of book, pages held distinct content, separate from a longer span of flowing text. While we cannot offer direct proof, it appears plausible that these extracts were printed for reference.

It is clear, then, when we examine behaviour aggregated across users, that ebooks are being used for reference and assigned reading purposes. This represents a digital behaviour that many have argued makes sense [20] or is likely to happen [59], but which has never before been positively shown in usage data.

4.9 Topic versus Keyword Clustering

Browsing at the physical shelves is something readers profess to value [17; 31; 44], and that arguably affects loan patterns. Browsing—or topic clustering—is a more contested issue online. It has proved inferior when compared to keyword search for focussed discovery [14], leading some to claim its possible redundancy. However, others say browsing is both important and under-researched [23; 41]. We analysed the loan histories, taking all co-borrowing sessions : those sessions when two books are checked out of a library at the same time (by the same reader)..

The distance between books in all 3017 co-borrowing sessions was calculated three different ways. First, we counted the nominal distance between co-borrowed books in a Dewey-sorted list of all books; this count represented topic distance. Second, a search was done using the title of the ebook to identify search neighbours. Third, we did pairwise searches of shared keywords between books in each co-borrowing session allowing us to maximise the chances of a search match. Both search calculations were on an index using metadata only, not full text. Co-borrowed books were only considered neighbours by any of these measures if they fell within ten books of each other consistent with our previous work in this area [38]. We then compared search distance and shelf or topic distance as explanations for co-borrowings: as shown in Table 4, search clearly explains more co-borrowings than topic (even within the low number of co-borrowings explained by any of these methods).

Table 4. Closest Neighbour, Co-Loans

Year	Search Only	Browse Only	Search Closest	Tie	Browse Closest
2013	88	70	78	22	46
2015	196	54	38	96	31

Clearly serendipitous discovery in this collection is low, echoing readers' worries about the rise of ebooks (see Section 2.3).

5. DISCUSSION

Copying was, in all contexts, surprisingly rare. One regularly stated concern of publishers is securing their content sometimes far more than they would be able to with print [20]. Others have complained of the negative impact of rights management on fair use [19] and prior research has demonstrated how DRM can frustrate users [18; 35; 54]. However, the low rate of copying seen by us suggests that literal reuse is in fact low, and DRM may be unnecessarily raising barriers to use.

The trade-off between online viewing and printing is complex. In any co-borrowing session, readers tended to focus either on printing encountered ebooks, or viewing them online. However, there was so sign of a clear division between readers who print and readers who view. Instead, individual readers used both strategies at different times. This reflects other work about individuals' book seeking behaviour, which suggests few users adhere strictly to a single strategy; the majority of users employ mixed approaches over time [42; 51].

It is interesting to note the high proportion of co-borrowings that are either exclusively or mostly print sessions, with little viewing;. This behaviour begins to look like print borrowing—triaging at the shelves and taking the books away for later, more extensive use [5; 17; 31]. While readers who printed did not (and indeed cannot, due to usage restrictions imposed by publishers) print entire books, there is similarly no indication that those who check out print academic books read them from cover to cover.

The regular use of printing indicates that digital books are often used as a delivery mechanism for print reading. This has been previously reported by Marshall [35], but there is no evidence of any shift away from print in 2015. The role of ebooks as print source material is probably going to endure for some time, despite many arguments that a new generation of students read exclusively online, and in spite of the associated individualized cost of printing[1] . Conversely, arguments that only print reading is deep or academic reading also do not hold true, with many readers displaying behaviours that look like deep reading over time in view mode, without printing.

We saw many types of loan activity. Similar to the JISC report on ebook reading [9], we saw many short loans with limited time-on-page; this looks like Adler et al's 'checking' reading or Marshall's triage reading [1; 36]. In contrast, there were a number of loans where the time-on-page was over a minute and a large number of pages read; these loans look very like the 'deep reading' as described by Adler et al. and Marshall. This deep reading could occur over multiple sessions; in contrast with print loans which (usually) persist for a number of days, each ebook loan lasts only 24 hours.

We examined two types of sequence: co-borrowing, and renewal-type sequences, where an individual reader returned to a book on a number of occasions. Neither of these are well addressed in the literature. Many quick co-borrowing sessions seem again to be in many cases a form of triage, particularly for print. As just noted, renewals may involve checking, but also longer-term multi-chapter reading, in contrast to the chapter-focussed printing and viewing that occurred in many popular documents. This style of

[1] See http://www.swinburne.edu.au/library/study-spaces-computers/print-copy-scan/ for print costs at Swinburne.

reading, probably driven by reading lists, is absent from Adler's categories, which are built on professional work.

Co-borrowing sessions came in two main forms: viewing-focussed, and print-focussed. There was minimal occurrence of printing in sessions that commenced with viewing, but not printing, a document; conversely there were a number of print-focussed sessions. Previous research has articulated the likely existence of 'print-to-read' strategies [35], and that seems a plausible explanation for the behaviours we observed here. However, elements of our data show other approaches, too. Some printing is of specific, atomic content such as tables or guidelines, which appears to be for reference purposes. Printing is often the last action in a series of loans of the same book, but it frequently has little overlap with the pages viewed online in a book. These facts suggest that this 'final' printing is not simply a method for preserving materials that have been viewed online for later offline reading. Printing and online reading appear to be complementary, rather than competing actions. Most users deploy both strategies in their reading, rather than turning exclusively to one or other.

Our data shows that, in contrast to recent studies of browsing in the physical library [38; 41], search appears a more likely explanation of co-borrowing of ebooks than traditional topic clustering. Indeed, compared to that data on co-borrowing in the physical library, rates of co-borrowing seem surprisingly low. Given that much of a library's reading occurs in the building, the print data almost certainly under-reports co-use. As noted by previous ebook studies [6; 35], the advantages of the print medium could explain part of this gap, the difference in discovery behaviour suggests that finding related ebooks is part of the challenge.

There has been a spate of recent research on how to provide browsing and non-search discovery for books online [23; 49; 60]; our findings underline the importance of this work. Search is not an analogue of, nor a replacement for browsing—not even faceted search [2; 25; 62], which is a search narrowing behaviour rather than a breadth strategy. We have known that search is not enough since the publication of early information seeking behaviour studies [3; 24; 34]. While some Similarly, given the dearth of topical co-borrowing in ebook collections topical similarity seems unlikely to entirely explain the value of browsing the library shelves, though a closer understanding of user cognition and approach is needed. Log-based analyses, as presented here, cannot address that gap in our scientific knowledge.

6. CONCLUSIONS

Logs have always had the potential to reveal new behaviours and validate users' self-reports. This potential for analysis of ebook data was noted in 2010 [35], however studies of ebook logs remain stubbornly limited in both number and scope. The majority of such studies address only that ebook borrowing takes place at all [7-9; 29]; the capacity to study user activity beyond simple borrowing has only been leveraged in our own earlier work [37; 39; 40] to the best of our knowledge. Readers' use of and attitudes to ebooks have been studied almost entirely using survey methods [15; 27; 46; 52] and experimental approaches [28; 32] limiting our capacity for understanding actual ebook use.

Our data reveals a range of user behaviours. Some of these behaviours, such as apparent deep reading and reference use, mimic print borrowing (in many cases to the point that readers are using ebooks as a print delivery service). Conversely some behaviours—such as the preponderance of singleton loans, and the heavy reliance of co-borrowing on search—appear to confirm the stated reservations readers have about ebooks, such as the loss of browsing and serendipitous discovery. 'Born digital' activities such as text copying are seen only on a limited basis, despite users claiming to value them. This means that born digital responses such as DRM are likely unduly heavy handed.

Overall, we now know that ebook borrowing is increasing in popularity, and includes both printing and viewing actions. Printing tends to involve a smaller proportion of each book than viewing; this is likely due to the interface support (or lack thereof) for each action. Similarly books that are borrowed together are more likely to share keywords than a Dewey classification; interface support for topic browsing in this collection is limited. Individual readers employ both printing and viewing in their interactions with ebooks, putting paid to both the argument that print is obsolete and that electronic reading simply doesn't happen—even individuals are not so strict about their preferences. Readers return regularly to books they have previously used, either examining new content and reading further, or returning to the same content for reference purposes. Some books are clearly being used as assigned reading, with the majority of content remaining untouched in favour of single pages or chapters. In summary, our analysis has revealed some facets of ebook use we might have suspected, some that mimic what we know about print use, some that are born digital, and some that are entirely surprising.

We have only grazed the tip of the iceberg with this analysis, however; the logs give us the capability to, for example, explore the impact of interface changes, to understand the relative value to users of different classification schemes, and to follow the behaviour of individual readers and books over much longer than a semester, to name a few. These and other interesting challenges remain future work.

7. ACKNOWLEDGMENTS

The authors thank Kim Tairi, Tony Davies and Justin Kelly at Swinburne University of Technology for access to the data that forms the basis of this paper.

During the writing of this paper, one of the authors said goodbye to Satchmo, one of the best cats who ever lived. They would like to thank him for 16 years of companionship, and dedicate this paper to his memory.

8. REFERENCES

[1] Adler, A., Gujar, A., Harrison, B.L., O'Hara, K., and Sellen, A., 1998. A diary study of work-related reading: design implications for digital reading devices. In *Proc. CHI 98* (Los Angeles, California, United States, 1998), ACM Press, 241-248. DOI= http://dx.doi.org/10.1145/274644.274679.

[2] Ballard, T. and Blaine, A., 2011. User search-limiting behavior in online catalogs: Comparing classic catalog use to search behavior in next-generation catalogs. *New Library World 112*, 5, 261-273.

[3] Bates, M.J., 2002. Toward an integrated model of information seeking and searching. *The New Review of Information Behaviour Research 3*, 1-15.

[4] Berg, S.A., Hoffmann, K., and Dawson, D., 2010. Not on the Same Page: Undergraduates' Information Retrieval in Electronic and Print Books. *J Acad Libr 36*, 6, 518-525. DOI= http://dx.doi.org/10.1016/j.acalib.2010.08.008.

[5] Blandford, A., Rimmer, J., and Warwick, C., 2006. Experiences of the Library in the Digital Age. In *Proc. CCCDT 06* (Tavros, Greece, 2006), Foundation of the Hellenic World.

[6] Brown, L., 2009. *Ebooks and the academic library: their usage and effect*, Masters Thesis, Aberystwyth University.

[7] Christianson, M., 2005. Patterns of use of electronic books. *Libr Collect Acquis 29*, 4, 351-363. DOI= http://dx.doi.org/10.1080/14649055.2005.10766084.

[8] Christianson, M. and Aucoin, M., 2005. Electronic or print books: Which are used? *Libr Collect Acquis 29*, 1, 71-81. DOI= http://dx.doi.org/10.1080/14649055.2005.10766034.

[9] CIBER, 2009. *Scholarly e-books usage and information seeking behaviour: a deep log analysis of MyiLibrary.*

[10] Dillon, A., 1992. Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics 35*, 10, 1297-1326. DOI= http://dx.doi.org/10.1080/00140139208967394.

[11] Hancock-Beaulieu, M., 1993. A comparative transaction log analysis of browsing and search formulation in online catalogues. *Program: electronic library and information systems 27*, 3, 269-280. DOI= http://dx.doi.org/10.1108/eb047145.

[12] Hancock-Beaulieu, M., 1993. Evaluating the impact of an online library catalogue on subject searching at the catalogue and at the shelves. *J Doc 46*, 4, 318-338. DOI= http://dx.doi.org/10.1108/eb026863.

[13] Hardy, G. and Davies, T., 2007. Letting the patrons choose: using EBL as a method for unmediated acquisition of ebook materials. In (Sydney, Australia, 2007), ALIA.

[14] Hearst, M.A. and Pedersen, J.O., 1996. Reexamining the cluster hypothesis: scatter/gather on retrieval results. In *Proc. SIGIR 96* (Zurich, Switzerland, 1996), ACM, 76-84. DOI= http://dx.doi.org/10.1145/243199.243216.

[15] Hernon, P., Hopper, R., Leach, M.R., Saunders, L.L., and Zhang, J., 2007. E-book Use by Students: Undergraduates in Economics, Literature, and Nursing. *J Acad Libr 33*, 1, 3-13. DOI= http://dx.doi.org/10.1016/j.acalib.2006.08.005.

[16] Hinze, A., Alqurashi, H., Vanderschantz, N., Timpany, C., and Alzahrani, S., 2014. Social Information Behaviour in Physical Libraries: Implications for the design of digital libraries. In *Proc. DL 14* (2014), IEEE, 107-116. DOI= http://dx.doi.org/10.1109/JCDL.2014.6970156.

[17] Hinze, A., McKay, D., Vanderschantz, N., Timpany, C., and Cunningham, S.J., 2012. Book selection behavior in the physical library: implications for ebook collections. In *Proc. JCDL '12* (Washington, DC, USA, 2012), ACM, 305-314. DOI= http://dx.doi.org/10.1145/2232817.2232874.

[18] Holt, K., 2010. E-book Sales Statistics from BISG Survey. In *Publishing Perspectives*.

[19] Houghton, S., 2011. Ebook Users' Bill of Rights. In *Librarian in Black*, San Rafael, CA.

[20] Hull, R. and Lennie, M., 2010. Why E-Textbooks Just Make Sense: An Academic and a Literary Agent Explain In *Publisher's Weekly Soapbox* PWxyz, New York NY.

[21] Jones, S., Cunningham, S.J., and McNab, R., 1998. An Analysis of Usage of a Digital Library. In *Proc. ECDL 98* (Heraklion, Crete, 1998), Springer, 261-277.

[22] Kantor, P.B., 1976. Availability analysis. *JASIST 27*, 5, 311-319. DOI= http://dx.doi.org/10.1002/asi.4630270507.

[23] Kleiner, E., Rädle, R., and Reiterer, H., 2013. Blended shelf: reality-based presentation and exploration of library collections. In *Proc. CHI 13* (Paris, France, 2013), ACM, 577-582. DOI= http://dx.doi.org/10.1145/2468356.2468458.

[24] Kuhlthau, C.C., 1999. Inside the Search Process: Information Seeking from the User's Perspective. *JASIST 42*, 5, 361-371. DOI= http://dx.doi.org/10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-#.

[25] Kules, B., Capra, R., Banta, M., and Sierra, T., 2009. What do exploratory searchers look at in a faceted search interface? In *Proc. JCDL 09* (Austin, TX, USA, 2009), ACM, 313-322. DOI= http://dx.doi.org/10.1145/1555400.1555452.

[26] Lau, E.P. and Goh, D.H.-L., 2006. In search of query patterns: A case study of a university OPAC. *Inform Process Manag 42*, 5, 1316-1329. DOI= http://dx.doi.org/10.1016/j.ipm.2006.02.003.

[27] Li, C., Poe, F., Potter, M., Quigley, B., and Wilson, J., 2011. *UC Libraries Academic e-Book Usage Survey.* University of California, Springer, California Digital Libraries.

[28] Liesaputra, V. and Witten, I.H., 2008. Seeking information in realistic books: a user study. In *Proc. JCDL 08* (Pittsburgh PA, 2008), ACM, 29-38. DOI= http://dx.doi.org/10.1145/1378889.1378896.

[29] Littman, J. and Connaway, L.S., 2004. A Circulation Analysis of Print Books and E-Books in an Academic Research Library. *Libr Resour Tech Serv 48*, 256-262.

[30] Losee, R.M., 1993. The relative shelf location of circulated books: A study of classification, users, and browsing. *Libr Resour Tech Serv 37*, 2, 197-209.

[31] Makri, S., Blandford, A., Gow, J., Rimmer, J., Warwick, C., and Buchanan, G., 2007. A library or just another information resource? A case study of users' mental models of traditional and digital libraries. *JASIST 58*, 3, 433-445. DOI= http://dx.doi.org/10.1002/asi.20510.

[32] Malama, C., Landoni, M., and Wilson, R., 2004. Fiction Electronic Books: A Usability Study. In *Proc. ECDL 04* (2004), Springer Berlin / Heidelberg, 69-79. DOI= http://dx.doi.org/10.1007/978-3-540-30230-8_7.

[33] Mangen, A. and Kuiken, D., 2014. Lost in an iPad: Narrative engagement on paper and tablet. *Scientific Study of Literature 4*, 2 (//), 150-177. DOI= http://dx.doi.org/10.1075/ssol.4.2.02man.

[34] Marchionini, G., Dwiggins, S., Katz, A., and Lin, X., 1993. Information Seeking in Full-Text End-User-Oriented Search Systems: The Roles of Domain and Search Expertise. *Libr Inform Sci Res 15*, 1, 35-69.

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

[36] Marshall, C.C. and Bly, S., 2005. Turning the page on navigation. In *Proc. JCDL 05* (Denver, CO, USA, 2005), ACM, 225-234. DOI= http://dx.doi.org/10.1145/1065385.1065438.

[37] McKay, D., 2011. A jump to the left (and then a step to the right): reading practices within academic ebooks. In *Proc.*

OzCHI 11 (Canberra, Australia, 2011), ACM, 202-210. DOI= http://dx.doi.org/10.1145/2071536.2071569.

[38] McKay, D., Buchanan, G., and Chang, S., 2015. Tyranny of Distance: Understanding Academic Library Browsing by Refining the Neighbour Effect. In *Proc. TPDL 15* (Poznan, Poland, 2015), Springer, 280-294. DOI= http://dx.doi.org/10.1007/978-3-319-24592-8_21.

[39] McKay, D., Buchanan, G., Vanderschantz, N., Timpany, C., Cunningham, S.J., and Hinze, A., 2012. Judging a book by its cover: interface elements that affect reader selection of ebooks. In *Proc. OzCHI 12* (Melbourne, Australia, 2012), ACM, 381-390. DOI= http://dx.doi.org/10.1145/2414536.2414597.

[40] McKay, D., Hinze, A., Heese, R., Vanderschantz, N., Timpany, C., and Cunningham, S.J., 2012. An Exploration of ebook Selection Behavior in Academic Library Collections. In *Proc. TPDL '12* (Paphos, Cyprus, 2012), Springer, 13-24. DOI= http://dx.doi.org/10.1007/978-3-642-33290-6_2.

[41] McKay, D., Smith, W., and Chang, S., 2014. Lend me some sugar: Borrowing rates of neighbouring books as evidence for browsing. In *Proc. DL 2014* (2014), 145-154. DOI= http://dx.doi.org/10.1109/JCDL.2014.6970161.

[42] McKay, D., Smith, W., and Chang, S., 2015. Down the Superhighway in a Single Tome: Examining the Impact of Book Format on Borrowing Interactions. In *Proc. OzCHI 15* (Parkville, VIC, Australia, 2015), ACM, 517-525. DOI= http://dx.doi.org/10.1145/2838739.2838766.

[43] Megarrity, L., 2010. Books Matter: The Place of Traditional Books in Tomorrow's Library. *ALJ 59*, 1/2, 6-11.

[44] Mikkonen, A. and Vakkari, P., 2012. Readers' search strategies for accessing books in public libraries. In (Nijmegen, The Netherlands, 2012), ACM, 214-223. DOI= http://dx.doi.org/10.1145/2362724.2362760.

[45] Nicholas, D., Huntington, P., Jamali, H.R., and Tenopir, C., 2006. Finding Information in (Very Large) Digital Libraries: A Deep Log Approach to Determining Differences in Use According to Method of Access. *J Acad Libr 32*, 2, 119-126. DOI= http://dx.doi.org/10.1016/j.acalib.2005.12.005.

[46] Nicholas, D., Rowlands, I., and Jamali, H.R., 2010. E-textbook use, information seeking behaviour and its impact: Case study business and management. *J Inf Sci 36*, 2 (April 1, 2010), 263-280. DOI= http://dx.doi.org/10.1177/0165551510363660.

[47] O'Hara, K. and Sellen, A., 1997. A comparison of reading paper and on-line documents. In *Proc. CHI 97* (Atlanta, Georgia, United States, 1997), ACM. DOI= http://dx.doi.org/10.1145/258549.258787.

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

[49] Pearce, J. and Chang, S., 2014. Exploration without Keywords: The Bookfish Case. In *Proc. OzCHI 2014* (Sydney, Australia, 2014), ACM, 76-79. DOI= http://dx.doi.org/10.1145/2686612.2686639.

[50] Pearson, J., Buchanan, G., and Thimbleby, H., 2010. HCI design principles for ereaders. In *Proc. Booksonline '10* (Toronto, ON, Canada, 2010), ACM. DOI= http://dx.doi.org/10.1145/1871854.1871860.

[51] Rowlands, I. and Nicholas, D., 2008. Understanding Information Behaviour: How Do Students and Faculty Find Books? *J Acad Libr 34*, 1 (1//), 3-15. DOI= http://dx.doi.org/10.1016/j.acalib.2007.11.005.

[52] Rowlands, I., Nicholas, D., Jamali, H.R., and Huntington, P., 2007. What do faculty and students really think about e-books? *Aslib Proceedings 59*, 6, 489-511. DOI= http://dx.doi.org/10.1108/00012530710839588.

[53] Saarinen, K. and Vakkari, P., 2013. A sign of a good book: readers' methods of accessing fiction in the public library. *J Doc. 69*, 5, 736-754. DOI= http://dx.doi.org/10.1108/JD-04-2012-0041.

[54] Shelburne, W.A., 2009. E-book usage in an academic library: User attitudes and behaviors. *Libr Collect Acquis 33*, 2-3, 59-72.

[55] Shibata, H. and Takano, K., 2014. Reading from paper versus reading from a touch-based tablet device in proofreading. In *Proc. DL 14* (London, United Kingdom, 2014), IEEE Press, 433-434.

[56] Stelmaszewska, H. and Blandford, A., 2004. From physical to digital: a case study of computer scientists' behaviour in physical libraries. *IJDL 4*, 2, 82-92. DOI= http://dx.doi.org/10.1007/s00799-003-0072-6.

[57] Takano, K., Shibata, H., and Omura, K., 2015. Effects of paper on cross-reference reading for multiple documents: Comparison of reading performances and processes between paper and computer displays. In *Proc. OzCHI '15* (Parkville, VIC, Australia, 2015), ACM, 497-505. DOI= http://dx.doi.org/10.1145/2838739.2838745.

[58] Takano, K., Shibata, H., Omura, K., Ichino, J., Hashiyama, T., and Tano, S.i., 2012. Do tablets really support discussion?: comparison between paper, tablet, and laptop PC used as discussion tools. In *Proc. OzCHI '12* (Melbourne, Australia, 2012), ACM, 562-571. DOI= http://dx.doi.org/10.1145/2414536.2414623.

[59] Thayer, A., Lee, C.P., Hwang, L.H., Sales, H., Sen, P., and Dalal, N., 2011. The imposition and superimposition of digital reading technology: the academic potential of e-readers. In *Proc. CHI 11* (Vancouver, BC, Canada, 2011), ACM. DOI= http://dx.doi.org/10.1145/1978942.1979375.

[60] Thudt, A., Hinrichs, U., and Carpendale, S., 2012. The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In *Proc. CHI 12* (Austin, Texas, USA, 2012), ACM, 1461-1470. DOI= http://dx.doi.org/10.1145/2207676.2208607.

[61] Wallace, P.M., 1993. How do patrons search the online catalog when no one's looking? Transaction log analysis and implications for bibliographic instruction and system design. *RQ 33*, 2, 239-253.

[62] Yee, K.-P., Swearingen, K., Li, K., and Hearst, M., 2003. Faceted metadata for image search and browsing. In *Proc. CHI 03* (Ft. Lauderdale, Florida, USA, 2003), ACM, 401-408. DOI= http://dx.doi.org/10.1145/642611.642681.