

# Lend me some sugar: Borrowing rates of neighbouring books as evidence for browsing

Dana McKay  
University of Melbourne  
Parkville, VIC 3010  
Australia

dmckay1@student.unimelb.edu.au

Wally Smith  
University of Melbourne  
Parkville, VIC 3010  
Australia

wsmith@unimelb.edu.au

Shanton Chang  
University of Melbourne  
Parkville, VIC 3010  
Australia

shanton.chang@unimelb.edu.au

## ABSTRACT

There is more to choosing a book than simply keyword searching. Browsing is a fundamental part of the information seeking process, and one that information seekers profess to value, though it has attracted little study. This dearth of research is undoubtedly in part because browsing is nebulous and difficult to quantify. In this paper we use a large circulation dataset from an academic library consortium to examine whether books in the library stacks are loaned in clusters, with a view firstly to confirming the existence of book browsing that has been reported anecdotally, and secondly to quantifying its impact on loan patterns.

## Categories and Subject Descriptors

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

H3.7. Digital Libraries: User issues;

## General Terms

Design, Human Factors

## Keywords

Libraries, digital libraries, classification systems, books, browsing, information seeking, log analysis

## 1. INTRODUCTION

Rowlands et al noted a surprising dearth of literature on book selection in 2007, an observation that remains largely true today [66]. What does it mean, though, to select a book? At the very least it must mean there are alternatives between which a reader can choose, many of which may meet that reader's needs. What we do not understand are the cognitive and environmental factors underpinning the selection of any specific book or books from among a range of available material—for example in a library.

We know from the literature on information seeking that browsing, serendipity, or being “open to new information” [37, 49, 56] are a natural part of the human information seeking process; we also know that this part of the process is poorly supported online [12]. For those readers who have clear criteria (for example the latest Booker Prize winner, or their university textbook) provided they have enough information [11] and the right vocabulary to search, finding what they need is simple [6]. For users with less well-defined information needs, though, the task is vastly more difficult [63].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

2014 IEEE/ACM Joint Conference on Digital Libraries (JCDL), September 8-12 2014, London, United Kingdom  
978-1-4799-5569-5/14/\$31.00 ©2014 IEEE

Information seekers have repeatedly demonstrated an appreciation of physical space for browsing: this is true of music shopping [21], book shopping [11] and in libraries [5, 31, 52, 71]. Work in this area has identified a range of benefits offered by physical space that are not commonly available in digital interfaces. Firstly, physical space allows users to begin their exploration without identifying any search terms (a task that may well be difficult [24]), or interacting with a search system (e.g. a catalogue [6]) that in all likelihood is unusable to them. Secondly, physical space allows users to switch back and forth between a close view (for example examining an index) to a broad view (looking at the other books on the shelf) very readily, a pattern of interaction that appears key to the book browsing experience [31].

We know browsing is important: it is a key part of information seeking behaviour [37], and maybe of intellectual endeavour [15]. Users like to do browse [71] and bemoan the lack of opportunity to do so online [31, 46, 52]. Given the recent explosion in ebook use [27], understanding browsing behaviour and its usefulness is more important now than ever. A good understanding of browsing is necessary to create good information seeking systems for ebooks; especially given that readers cite the lack of browsing opportunities such as those provided by library shelves as a reason to avoid ebooks [31].

Why, given its importance, is there so little literature on browsing behaviour? We recognise that there are a handful of studies of online browsing systems (for example [7, 55, 63]), and that some studies have investigated behaviour at the library shelves (for example [31, 64]), but there is nothing that attempts to measure or quantify browsing. This dearth is possibly because quantifying serendipity is so difficult [30], and the success of any browsing system is necessarily hard to measure.

This paper uses a large data set recently released by the OCLC<sup>1</sup> (Online Computer Library Center) to attempt to quantify the browsing that occurs among the library shelves. By examining the dates on which books were checked out and their location on the shelf, any spatio-temporal clustering of loans should become apparent. Clustering of this kind can be assumed to be browsing, given both the repeated user reports of its value [5, 31, 36, 46, 52, 67, 71] and early studies observing its frequency [29, 43]. We discuss this further in Section 2.5.

There is a history in digital library research of using user behaviour in physical libraries to inform digital design, for example [20, 21]. Given that library shelves are far from bastions of usability [53] and cannot be rearranged to meet users' needs, any evidence of browsing in a physical library environment should be taken as strong evidence to support the development of effective mechanisms for online book browsing.

<sup>1</sup><http://www.oclc.org/research/activities/ohiolink/circulation.html?urlm=159699>

The remainder of this paper is broken up into five sections: the literature background for this work, the methodology used to quantify browsing, the results of this study, a discussion of those results, and conclusions and future work.

## 2. BACKGROUND

To talk about browsing, we must first define it: a definition grounded in browsing literature is presented in Section 2.1. Section 2.2 describes how this behaviour fits into the human information seeking process; Section 2.3 looks at what we know about information seeking in libraries; Section 2.4 examines what we already know about book selection practices, and Section 2.5 makes the case for clustering of loans representing browsing.

### 2.1 On Browsing

Browsing can be described as undirected information seeking allowing for serendipitous discovery of useful resources [15]. Perhaps the seminal work on browsing in Information Retrieval (IR) is Bates 1993 paper [2], which discusses a ‘berrypicking’ model of information retrieval, in which both users’ thinking and search behaviour are influenced by the documents they encounter as part of the process. She notes that online browsing is key to facilitating this interaction between seekers and information. Bates furthered this work with a literature-based definition of browsing in 2007, where she noted that browsing includes visual glimpsing of ‘a scene’, noting items of interest, honing in on them and evaluating them. If more than one item was identified, in her model they are evaluated sequentially [3].

Despite the need for online browsing being noted in 1993, there is a dearth of work in this field. A 2004 paper presents three approaches in Greenstone digital library software, though none of the systems presented was more than proof of concept [55]. More recently Pearce et al. have developed an ‘exploration engine’ [63] and found information seekers to be very successful with it when performing browsing-suited information tasks. In the past two years two systems have been developed to facilitate browsing books in libraries [36, 74]. These systems present books in a range of ways, including a shelf-like display, groups by size or colour, but testing (while looking initially positive) has been limited.

Some might argue that search result presentations facilitate information seekers’ browsing, but we know the vast majority of users view only ten search results [70] on a page, or one or two results in any depth [28]. These small numbers are hardly the openness to serendipitous discovery described above as being part of browsing. Similarly online recommender systems such as Amazon’s ‘customers who bought ... also bought’ feature may be considered browsing, however these systems require prior knowledge of the user’s purchasing habits or priming with ratings [30], thus not allowing the full scope of browsing offered by the library shelves (moreover shelf browsing is still more popular among many user groups [30]).

### 2.2 Human Information Seeking Behaviour

There are a number of models of human information seeking behaviour, with the two most fundamental perhaps being Marchionini [49] and Kuhlthau [37]. These models are both broadly linear, and include discovering an information need, formulating a question, posing a question, examining the answers, refining the question, synthesizing the answers and deeming the need met. In both models browsing appears as a key component used for refining information needs or corpus understanding. In both models browsing is closely interleaved with search, indeed information seekers move between the two seamlessly.

The classical models of information seeking behaviour, though, do not take into account either the social dimensions of information seeking, nor the serendipitous discovery of information the user did not know they wanted or needed. Contrast these with McKenzie’s [56] model, which is rooted in the social information seeking often seen in observational or ethnographic studies (for example [10, 16, 62, 73]) shows information seekers “putting themselves in the way” of information, whether intentionally or not.

Taken together, these exploratory behaviours—putting oneself in the way of information and examining information in the environment—constitute exploration or browsing. Online facilities for browsing were noted to be largely absent and poor at best in 1993 [2]; this problem still exists today [12].

Contrast the online environment with the physical environment, where users report browsing in a range of information seeking tasks: music shopping [21], book shopping [11] and visiting public [62] and academic [71] libraries to name a few.

### 2.3 Information Seeking In Libraries

Information seekers’ interactions with libraries can be taken as interactions with complex information systems [15, 46]. These systems are made up of people (staff and other library users) technology (including catalogue and search systems), information (including but not limited to books) and the physical environment (the layout of space and the shelves).

Interactions with library technology are well-studied; transaction log analyses of catalogues abound (see for example [39, 57, 75]). These analyses typically find catalogues wanting, as do usability studies [6, 22, 26]. Lack of browsing capacity is a frequently mentioned drawback [46, 52], as are opaque query formulation mechanisms [6] and poor user interfaces [69]. Beaulieu noted that catalogues support browsing poorly, but that information seekers were more successful when they began with browsing, perhaps suggesting one of the major flaws in these systems [28].

Interactions with library staff often occur as a result of or in place of failed interactions with other parts of the library system; users speak to staff when they are having problems with the shelves or catalogue [32, 53], or simply do not know where to start [60]. There have been a number of studies of users’ interactions with library staff (see for example [9, 16, 60]); these interactions are especially useful to users who have complex information needs, or who have reached the limits of their library understanding

Library shelves are carefully designed to encourage serendipity [8, 15], indeed Svenonius notes, they are unusual because document addresses (i.e. call numbers) contain information about the semantic content of the document [72]; as we shall see below, readers find this valuable for browsing. Shelf browsing is an example of McKenzie’s ‘putting oneself in the way of useful information’ [56] activity (also seen in work on serendipity [15, 23, 45-47]). This paper attempts to address how successful that arrangement is in affecting borrowing behaviour using circulation data.

Given what we know about how users interact with the component parts of a library system, how do they go about choosing books?

### 2.4 Book Selection Literature

Little work has addressed book selection in any environment beyond shelf location, however it is clear that search is not the whole answer to choosing books. A 2008 study by Rowlands and Nicholas shows academic library users employing a range of book

selection strategies, from internet searching through library searching and browsing, to asking a friend [65]. A study of queries asked in bookshops in the UK [11] demonstrates that shoppers often lacked even the most basic vocabulary to describe their needs in a way an information system could interpret. A study done across a number of libraries in New Zealand and Germany [31] showed users interacting with books in a lightweight way during their decision making process—exploring more deeply as they discovered interesting things without losing context. This backgrounding of context is not readily available online, yet this behaviour directly reflects the close intertwining of search and exploration seen in the classical information seeking models and Bates’ definition of browsing [3]. A further study identifies social elements to book selection that are well supported by physical shelves [18]. There has been a single attempt to emulate shelves online [7], but it was directed at children and did not provide the context-rich experience browsing in a physical space offers. Even this limited approach, though, was enjoyed by children and made them more effective information seekers.

Part of book selection, as alluded to above, is examining individual books more closely—browsing within. Work on within-book browsing has shown us that within books the cover, index and table of contents are relevant for decision making [4, 48, 50, 51], and that users appreciate a naturalistic experience when selecting among online books [41]. The simplicity of examining a book in print is part of what facilitates good browsing—it is easy to move on to the next thing of interest.

Physical spaces appear to be much better suited to browsing—the serendipitous discovery of information in a semi-directed way—than any online systems created thus far [46, 47]. This finding is borne out in the literature, at least for some user groups [46, 65]. This appears to be because physical co-location creates context and allows information seekers to put themselves in the way of useful information [23, 47]; certainly information seekers believe this to be useful.

## 2.5 The case for browsing

Loizides’ document triage model [42] shows that at each stage of selecting online documents, the choices become fewer: information seekers narrow down search results and then examine a few documents from the chosen results set. This narrowing is disrupted by visiting the shelves in the process of selecting a book: information seekers are necessarily exposed to the shelves.

The shelves are (as Kleiner et al. clearly point out [36]) an ideal fit for human browsing as described by Bates in 2007 [3]: lots of visual information and the ability to easily investigate objects of interest sequentially. In light of this fit, it is hardly surprising that in library literature it is considered obvious that shelf browsing is an important part of the information seeking process [68]. We do not have to take this on faith, however. Users in a number of studies have reported the value of browsing in their book selection process, in both academic [31, 46, 65, 71] and public libraries [62, 67]. This value comes from the implicit or explicit understanding that libraries arrange the books so that serendipitous discovery might happen [46], and from the availability of cues (such as dust on books, the number of copies, the covers) that are simply not available online [46, 71]. Indeed, 55% of users in a study from last year said that shelf browsing is still an important part of information seeking [36].

In addition to readers’ self-reports, observational studies have identified browsing incidentally in studies of children seeking fiction [19, 64] (though not nonfiction [58]). Adults have been observed shelf browsing both fiction [68] and non-fiction [31, 71]

in naturalistic studies, though sample sizes are necessarily limited. Finally, a 1993 study [43] based on the circulation statistics of a small group of users and a single Library of Congress Subject Heading also identified books being borrowed in groups, and called it browsing.

We can see browsing takes place, and users say they value it, but it also appears to impact borrowing patterns. In the only study on browsing and borrowing to date to date, more than 50% of readers who found a book they had identified in the catalogue also borrowed at least one further book that had not been previously identified in search results [28].

Contrast this significant evidence with the case for search result browsing. At the time when the data was collected (2007-2008) catalogue usability was poor [38, 69], users did not really understand why they got the search results they did [46] and tools that may have helped users browse (such as faceted search [38]) were still not in widespread use [69]. The number of search results viewed in catalogue searches is small [28], and work contemporaneous with the data collection [46, 65] suggests that users were familiar with the physical layouts of their libraries and considered the library shelves a better way to find books than the library catalogues.

Given this preponderance of evidence for browsing, it seems highly likely that any clustering of loans of physical items is likely to be as a result of shelf browsing, rather than any other behaviour.

What is lacking in the literature outlined above, though, is any large scale attempt to quantify the impact of browsing on loan patterns. The large data set used in this study makes it possible to essentially perform a transaction log analysis on physical behaviour. This study attempts use that log analysis to demonstrate the impact of co-locating like items in a physical space (and thus likely facilitating browsing) with a view to understanding just how important browsing is for online book selection.

## 3. METHODOLOGY

Transaction log analysis is a long-standing method in digital libraries (see for example [34, 44]); it can be applied to behaviour in a range of online systems, including library catalogues [39], web-scale search [57], journal databases [59], and internet search engines [70]. In this study, though, while the data exists online the technique is applied to interactions in physical space. Transaction log analyses are an excellent approach to understanding what users do when unobserved, however user motivations, satisfaction and non-system interactions cannot be imputed; the same is true in this study.

This section first describes the data set on which we performed our transaction log analysis, then describes the library selection method and the pre-processing necessary to use the data. Finally we describe the tests used to measure the impact of browsing.

### 3.1 The OCLC Circulation Dataset

In 2013 the OCLC released a very large data set, comprising circulation records for print books (not ebooks or journals) from the libraries in its OhioLink consortium. These records span a year over 2007 and 2008 [61].

The institutions in OhioLink are—with one exception—academic libraries. Most are open-stack (meaning the books can be browsed and checked out by library members) with a few exceptions; some libraries, such as OSU, have closed-stack storage.

For each book the circulation data includes its unique item number, which institution held it, which library and collection that book was found in, whether the book may be loaned, how many times it was loaned during the study period and the date on which it was most recently loaned. Having only the most recent circulation date means that we cannot always know whether two books were loaned on the same date, the data required careful analysis to resolve this. Bibliographic information, including call number and title, is stored in a separate MARC (MACHine Readable Cataloguing) file.

The OCLC have already calculated a loan rate for all collections [57], and even started some institutional analysis, however their approach has been limited to investigating broader circulation patterns, rather than looking at the specific issue of browsing.

### 3.2 Data Selection and Processing

Analysing browsing on the basis of circulation data required significant pre-processing. Initially the classification scheme used by each library in the OhioLink Consortium was identified by searching each library's catalogue for 'Shakespeare's Complete Works', or, if that returned no results, 'The Holy Bible' (these works were selected on the basis that most American libraries will have at least one of them). Reading comprehension SAT entry requirements were then checked for all four-year and graduate level colleges and universities. The SAT metric was used to select those institutions at which users were most comfortable using the library [33] and to ensure as far as possible similar students at each institution. For any institutions with more than one library circulation data from the main library was selected. Six sample institutions (three using Dewey Decimal Classification and three using Library of Congress) were selected on the basis of SAT reading scores required for entry. The Library of Congress (LC) libraries chosen were Ohio State's Thompson Library (OSU), Case Western Reserve's Kelvin Smith Library (CWRU) Cedarville University's library (Cedarville). The vast majority of books in each of these libraries were shelved using LC, though each of them had a fiction collection about which shelving information was not readily available (these books were therefore excluded from this study). The Dewey Decimal Classification (DDC) libraries selected were Oberlin's Main Library (OberlinDDC), Denison's library (Denison) and Ohio Northern's library (ON). Like the LC libraries, the Dewey libraries had fiction collections that were not shelved with the main collection, however each also had significant proportions of their collection shelved in classification schemes other than Dewey. The most notable of these is Oberlin's Main Library, which is partly Dewey and partly LC in non-contiguous shelving sections. To provide a better comparison between Dewey and LC, then, Oberlin's LC shelves (OberlinLC) were also extracted and analysed.

Once libraries were selected, call numbers and titles for each item were extracted from classification data and matched to circulation data to form the collections listed above. Summary statistics were then generated for each collection: number of books, earliest and most recent circulation date, overall circulation rate and annual circulation rate. Collections were then sorted by call number to mimic the library shelves, so that the impact of browsing could be identified. To test for this impact, two comparisons were performed for each collection.

### 3.3 Testing For a Neighbour Effect

The first comparison used a random sample of loaned books, determining for each book how many of its ten nearest neighbours (five in each direction on the shelf) had been loaned on the same date as its most recent circulation date. Five books in each

direction seemed the best initial compromise between too narrow and too wide—it gives a total scan length of 11 books, which can be reasonably assumed to be relatively closely co-located without necessarily being in the same call number range. Books with near neighbours that had been borrowed after them were excluded, as there is no way to tell that those books were not also loaned on the same day as the index book (because only the most recent loan date from each book is recorded). The ten neighbouring books were then examined to see how many of them had been loaned on the day before the date on which the index book was loaned. While it is impossible to guarantee that the index book was not also loaned on this date, library processing times make it extremely unlikely. This approach allows us to compare loans in a paired set of books with and without a neighbouring loan. By making this comparison, we are able to ensure that books from the same area of the shelves are not loaned simply because the subject area is popular on a given day (e.g. textbooks at the beginning of term); if there is a neighbour effect present here the concurrent loans cannot be simply ascribed to topic popularity.

The second comparison used the books loaned on the final or penultimate date of data collection for each collection. For three collections in two libraries the penultimate day was used as the loan rate on the last day was small enough as to suggest only renewals occurred on that date. Books being borrowed on the final date for which circulation data was recorded can have a true loan rate calculated for their neighbours (again, five in either direction), because the circulation dates associated with these books cannot have been overwritten by a more recent loan. These ten-book samples were then compared to 1000 random ten-book samples from the same collection to determine the comparative loan rate for the final circulation date. This comparison accounts for a particular loan date (e.g. near the end of term) being very popular and shows the specific influence of having a nearby book loaned.

The limitations of the data—in particular recording only the most recent circulation date—required careful consideration when planning these tests. This limitation, in fact, was the driver for performing two tests: by pairing books at a single location we can control for location effects, but are subject to inaccuracies of date data. Conversely by testing only books loaned on the last day against random books, we risk some effects based on material topic, but none on the basis of date. If both tests point to a neighbour effect, however, we can reasonably assume that the effect exists and browsing is the likely cause.

Finally collections were compared to see if there were any major differences between libraries, and classification schemes were compared to determine whether one could be said to be superior to the other for subject browsing.

## 4. RESULTS

The results section is broken down into five parts: Firstly we provide summary statistics for each collection in the study, then we examine whether browsing can be demonstrated by looking at loan patterns. Next we examine the size of any browsing impact, then we make a comparison between DDC and LC libraries. Finally we address the limitations of this study and provide a summary of our findings.

### 4.1 About the Collections

In this section we provide summary statistics about each collection used in the study to give some background for our analysis.

As shown in Table 1 overleaf, annual circulation rates (the number of books that are checked out at least once) for 2007-08 vary between 8 and 18%, and are broadly reflected in total circulation rates (the percentage of books that has been taken out at least once over the entire time data has been collected for each library).

Clearly, then, the vast probability is that any given book will not be checked out even once in a given year, much less on a given day. The small numbers seen in this study necessarily mean that the size of any effect appears small.

**Table 1: Summary statistics about test libraries**

	Lib type	First circ date DDMMYY	Last circ date DDMMYY	Number of books	Overall circ rate	Annual circ rate 07-08
OSU	LC	31-12-69	06-05-08	462328	0.65	0.18
CWRU	LC	21-09-87	14-04-08	497763	0.50	0.12
Cedarville	LC	27-12-93	06-05-08	124260	0.74	0.17
Oberlin LC	LC	31-12-69	17-04-08	5675663	0.41	0.09
Oberlin DDC	DDC	31-12-69	17-04-08	162174	0.45	0.11
Denison	DDC	18-02-93	14-04-08	103691	0.56	0.16
ON	DDC	16-03-93	22-04-08	76972	0.47	0.08

## 4.2 Does The Data Show Browsing?

As mentioned in Section 3, we used two tests to check for browsing: the first used randomly selected books and dates, and checked the same book set on days when an index book had been loaned and when it had not been. The second test compared books adjacent to those loaned on the final date in each collection's circulation records with books adjacent to randomly selected books and noted which books in each adjacent set had been loaned on the final day.

**Table 2: Loans of the same book on with or without a neighbouring loan**

	Neighbouring loans	No Neighbour	Significance	$\chi^2$	df
OSU	304	14	p<0.0001	264.779	1
CWRU	402	21	p<0.0001	345.020	1
Cedarville	479	49	p<0.0001	353.225	1
Oberlin LC	313	26	p<0.0001	243.349	1
Oberlin DDC	177	12	p<0.0001	143.065	1
Denison	161	17	p<0.0001	115.396	1
ON	306	21	p<0.0001	248.687	1

### 4.2.1 The Same Books on Two Dates

This section will outline the results of pairing books and testing on a different day. For each collection, 2000 books were used as index books, meaning 20000 books were examined on each of

two possible loan dates. In the table below 'Neighbouring loans' refers to the number of loans in that 20000 set on the date on which the index book was loaned; 'No Neighbour' is the number of loans on the day prior to the most recent loan of the index book for each set. To test the significance of results, we used a two-tailed chi-squared test with Yates' correction.

As we can see from Table 2 above, every collection in this set generates a statistically significant result at the 0.1% confidence interval for this test, meaning that a book being in the five books either side of a book on the date that it is loaned is much more likely to be loaned itself on that day than on the day before. The confounding factor here is that we cannot (with the limitations of this data) prove that the index book or any of the other books around it were not also borrowed on the day before. To address this we performed a second test, looking only at the last date on which loans were recorded for any book in the collection.

### 4.2.2 Books Loaned on the Last Day of Circulation

The second test we used took as a control set books loaned on the final or penultimate day on which circulation data was recorded for each collection. The libraries marked with an asterisk use data from the penultimate day because the final day sets were so small as to be most likely renewals on a day when the library was closed.

**Table 3: Impact of neighbouring loans on final day loan rates**

	Last day loans	Neighbouring loans (/ last day loans *10)	Random loans (/10000 books)	Significance	$\chi^2$	Df
OSU*	112	8	0	p<0.0001	61.890	1
CWRU	91	16	3	p<0.0001	133.534	1
Cedarville	77	8	6	p<0.0001	45.538	1
Oberlin LC*	136	50	3	p<0.0001	334.991	1
Oberlin DDC*	55	40	1	p<0.0001	1196.062	1
Denison	92	22	14	p<0.0001	123.192	1
ON	29	10	2	p<0.0001	77.923	1

Because we definitively know that no books were borrowed after this date we can know for certain for any set of ten books how many of them were loaned on the final date; no more recent circulation date could have overwritten the dates for those books. As such we counted the number of books loaned on the final loan date in each set of ten for each final loan date index book. We then compared this with the ten books around 1000 random books—i.e. 10000 books—in each collection to see whether we could confirm an effect. Results are shown in Table 3: last-day books is the number of books loaned on the final date collection date for each collection; neighbouring loans is the total number of books loaned in the ten books either side of a book loaned on the final circulation date; random loans is the number of loans on the final circulation date in a set of ten books around a randomly selected index book. Statistical significance, as for the last test, is

\* These collections used data from the day before the final date for which circulation information appears in the data as the final day sets were small and likely represented only renewals.

calculated using a two-tailed chi-squared test with Yates' correction.

Again, this test shows statistical significance for all collections. What this test cannot address is any local effects: whether, for example, books from a certain part of the collection are under heavy use on a given day because an assignment is due. The previous test, however, did address this issue and still found significance. Between the two tests, then, it is reasonable to assume that the effects seen are directly related to a recently-loaned neighbour. The logical explanation for this is shelf-browsing, i.e. users encountering neighbouring material that is also relevant to their information needs as they collect a target book; or indeed deliberately seeking books from a shelf-location known to be a good source.

### 4.3 Quantifying the Impact of Browsing

In the previous sub-section we demonstrated that browsing is, in fact, evident in the data about book use—that a book is statistically more likely to be borrowed if it is near another book that is borrowed on the same day. The size of that impact is the question for this section.

Our initial approach to determining the size of the effect was to calculate the  $\phi$ -coefficient for each test, however in all but one case (where it came out to 0.33) these were small, in the range 0.06-0.17, with most around 0.11. This seems at first to indicate a small effect, but when we consider the tiny proportion of books that are loaned on any given day, effects will necessarily be small.

In light of this we turn to descriptive statistics. Looking at the increase in loans as a proportion shows that having a near neighbour borrowed increases the likelihood of any given book being borrowed by at least ten-fold in all but two tests.

The final date collection statistic compares the likelihood of a book being loaned on the final circulation date in when it was in the ten books nearest a book loaned on the final date with the overall collection circulation rate (number of books loaned on the final date/number of books in the collection). Results for all three comparisons are shown in Table 4 below.

**Table 4: Size of the neighbour effect calculated three ways**

	Paired dates	Final date test	Final date collection
OSU	21.71	Undefined	294.85
CWRU	19.14	58.61	961.745
Cedarville	9.78	17.31	167.66
Oberlin LC	12.04	122.55	15342.95
Oberlin DDC	14.75	727.27	2144.45
Denison	9.47	17.08	269.52
ON	14.57	172.41	915.24

The population-based statistics show the most dramatic increases in odds; this is simply due to the small sample size and large population size involved. Nevertheless, even pairing books, a nearby loan increases loan likelihood by close to an order of magnitude in every case, if not more. This data demonstrates that browsing has a significant local impact, affecting a loaned book's nearest neighbours, but that this impact is nearly undetectable in overall collection circulation patterns due to low circulation generally.

### 4.4 Dewey vs. Library of Congress

It is apparent that browsing happens, and that its impact is significant, but is one classification system better than the other?

By comparing the proportion of neighbour books that were borrowed on the final recorded circulation date between DDC and LC collections we can begin to understand which classification scheme might better support browsing. The table below shows loaned and unloaned books in neighbour sets in Dewey libraries in comparison with LC libraries. To ensure any significance was not due to of some undetectable difference between users or library staff, we also show statistics for the two types of books within Oberlin library.

**Table 5: Neighbour effect in Dewey vs LC**

	DDC		LC		Significance	$\chi^2$	Df
	Loaned neighbours	Unloaned neighbours	Loaned neighbours	Unloaned neighbours			
All libraries	71	1689	82	4078	p<0.0001	20.094	1
Oberlin only	40	510	50	1310	p=0.0012	10.94	1
All but Oberlin	32	1778	32	2768	p=0.1005	2.698	1

Using the chi-squared test again we see that there is some difference between DDC and LC for browsing, but that this is mostly seen in Oberlin Library. Given that Oberlin does not have the full range of books in either classification scheme, it is possible that this apparent difference is due to subject matter, an interesting finding in and of itself. We know that readers are more or less likely to use ebooks based on discipline [14], so it would not be surprising to find browsing behaviour differences as well.

### 4.5 Limitations of This Study

This study will almost certainly underestimate the neighbour effect. There are a number of reasons for this: Firstly not being privy to shelf layout that we cannot accurately determine which books were above and below a borrowed book, only those on each side. Secondly the data does not distinguish between renewals (which may be significant proportion of loans) and new loans. Thirdly we cannot tell from the data exactly which books were available on the shelf on the sample date—loan periods may have changed (and are not available for the 2007-2008 period of this study), and books may be missing or off the shelves for repair or in use within the library. Finally, absent data at the patron level, this study cannot address the type of serendipitous discovery seen in [31, 67] in locations more distant than very nearby books. Any neighbour effect detected is therefore likely to underestimate actual browsing.

A possible confounding factor with this data is that shelves are arranged by topic. It is possible, though unlikely, that the neighbour effect represents books discovered by individual patrons in repeated catalogue searches. Given the evidence for browsing discussed in Section 2.5 though, this seems unlikely.

The topic-based classification scheme of all the shelves also means this study offers no understanding of the usefulness of other metadata types for book browsing, usefulness for which there is anecdotal evidence in other studies, for example [11, 65].

This study does not measure the amplitude of the neighbour effect at different distances, a task that remains for future work; nor does it account for duplicate books (which may affect the outcome but

are a tiny proportion—fewer than 0.8%—for libraries where this was checked).

Finally, this study includes only academic libraries; where the behaviour may be quite different to public or specialist libraries (see for example [31]); the findings therefore are not generalizable to these types of libraries.

## 4.6 Summary of Results

In this section we have established that there is a clear and significant ‘neighbour effect’ in book loans, where a nearby neighbour being loaned greatly increases the probability of any book being loaned itself. This clustering, given that it happens when you compare books by location and by date, almost certainly represents browsing in the form of interaction with physical shelves to find similar items for a broad information need. True, we cannot prove that these books were not all borrowed at different times of day by different people, but the simplest explanation by far is that some browsing is taking place. In the following section we will address the interaction between this newfound understanding and digital libraries research.

## 5. DISCUSSION

In this section we discuss our findings in light of the work on digital libraries. Firstly we compare our findings to what we already know about browsing in an online space, secondly we examine the applications of our findings to the online space and thirdly a discussion of this study in light of what it does not tell us.

### 5.1 Online Browsing and Search

An obvious question to ask, given that we have been able to quantify browsing for books in a physical library is what relationship does this bear to online browsing?

As mentioned in Section 2 there is very little literature on browsing online, however thinking about the process of borrowing a book from a library—searching the catalogue, then locating it on the shelves (if it is present in the library [71]), then checking the book out—the most common analogue is probably looking at search results. In information retrieval, clicking on more than a single result is seen as “noise” rather than user evaluation of possibilities [17]. The raw number of clicks per query is somewhat variable, but seems to hover between 1 and 2, suggesting little browsing of search results [1, 35, 76].

When looking at search result behaviour more broadly we see that what browsing does occur occurs within the first page of results. Jansen and Spink’s work from 2001 [70] shows that the majority of people look only at the first page of search results and no further—viewing ten result snippets, though not ten results. This is a very small pool for browsing, in comparison with what information seekers will see in a library, and offers little opportunity for serendipitous encounter [3, 45].

Looking closer to academic libraries, transaction log analyses have not typically looked at how many search results users click on, a surprising omission in an otherwise well-studied area. The exception to this is Beaulieu’s 1993 study, which showed the vast majority of users examining only a single result in any depth, and of those who examined more an even more significant majority only viewed two [28]. A large study of journal article access that compared access strategies—search and various types of browsing—shows that those who begin with search access fewer documents [59], though not how many documents they access. This reduced breadth of access is further evidence for the goal-driven behaviour seen in search.

It is unsurprising in light of earlier work [31, 54] that search result browsing, including the in-depth sequential examination of items suggested by Bates’ definition of browsing, is not seen online; search result lists do not provide the same easy interactions as physical shelves. It is also clear that information seekers who use search also use library shelves as an information source, thus search is not providing everything they need [65]. Indeed, even those users who browse search results go on to browse the shelves as well [27], showing that both activities have a place in book selection.

This study does not dispute the place of search and search result browsing in book selection, though we do not view users selecting more than a single book as noise. What this study does offer is the beginnings of a precision/recall like metric for browsing books in a physical space, and the numerical probability that it is important.

It appears that it may be slightly more common for users to check out more than one book than to view more than one search result, further confirming the value of the shelves for serendipitous discovery. Increasingly, though, it should be possible to provide exploration services online that are as good as or better than those in the physical world. Avenues for this are discussed below.

### 5.2 Implications for Online Browsing

There is a long history in digital library research of using studies of human information seeking behaviour in physical environments to inform digital interface design, see for example [19-21].

Information seekers have told us that they value the browsing experience they get in physical libraries [5, 71]; it is also likely that the serendipitous discovery offered by these spaces is both used [43] and beneficial to academic endeavour [15, 45]. We also know that browsing supports more successful information seeking in libraries than search alone [29]. Similarly, users do not get the same experience from recommender systems as they do from shelf browsing, because recommender systems have to be primed with knowledge about users and are not always accurate [28]. What is widely acknowledged, however, is how limited physical browsing in libraries is [8]: some libraries have tried as a result to do away with classification schemes altogether [13, 25].

Nevertheless, information seekers do browse by subject, especially in academic libraries [28, 53]. In a 1993 observational study over 50% of information seekers who successfully located a book in the catalogue also selected at least one more book from the shelves [29]. Similarly a small analysis of 96 users also shows them selecting multiple books from the same areas of the shelves [43], and that books in one very narrow subject heading tended to be selected together. The study presented in this paper unequivocally demonstrates that physical browsing even in this most limited of capacities occurs on a broad scale even when users are unobserved.

Online browsing has the potential, however, to offer so much more than can be offered by the physical shelves. Users have all sorts of nebulous, ill-defined needs for books including the colour of the cover [11, 18, 19], recency of usage [62, 71], book length [74], themes [19, 62] and print size and number of images [31, 64]. These ill-defined needs are exactly the kind best met by browsing [37]. The broad range of ways that books might be classified and the contextual nature of how information seekers might wish to browse is further expanded on in [8]; suffice it to say that browsing needs are widely varied. In the physical world, shelf arrangement is fixed; online we have the option to rearrange the shelves according to user needs.

There have been some attempts to provide exactly this type of exploration online [55, 63], however they have not, thus far, been applied to books except for one notable exception. Borgman offered a shelf-based metaphor within a digital interface to children in 1995 [7], and despite its low fidelity and limited customisability it was well received.

With the massive uptake in ebook use in both recreational [27] and academic arenas [40], and the common lament of losing browsing in the move to ebooks [31, 52], our research both provides a measure for successful browsing—the neighbour effect should be close to that in physical space—and reconfirms the need for it.

### 5.3 Browsing Behaviour

Like traditional transaction log analysis, this study leaves much about the actual user behaviour that occurs behind the numbers opaque. Just as an ordinary transaction log analysis can only track what users click on, this one only identifies what they borrow. We know that readers examine many more books than they borrow [31], and that there is a perceived cost to borrowing a book [52]. As such while this study strongly suggests the existence of browsing, it gives little insight into its nature

Similarly, this study provides no insight into within-book browsing an important part of book selection behaviour [31]. To support book selection online, we need to better understand these behaviours, perhaps starting with log analyses of online book selection interfaces.

This study has reconfirmed what users have self-reported about academic library shelves: that visiting the shelves is a valuable experience (see Section 2.5); this study even quantifies that value. What is not clear, however, is what frustrations, if any, users have with the shelf-browsing experience, and how useful or important the other kinds of browsing mentioned in Section 5.2 are. Again, to develop truly effective and usable browsing and book selection systems, we need to have a broader understanding than this study can provide.

Finally, in all models of information seeking searching and browsing are ideally interleaved [37, 49]; in the physical library they have necessarily been separated traditionally (though smartphones are changing this). This study, not having search data, does not address the interaction between searching and browsing.

The best approach to understanding book browsing beyond this study (which strongly suggests it happens and quantifies it) is to begin with observing it, both in person and online. This observation, and whatever further study it could give rise to, are avenues for future work.

## 6. CONCLUSIONS AND FUTURE WORK

In this section we will firstly draw conclusions from this study, then examine avenues for future work.

### 6.1 Conclusions

In this paper we have presented a quantitative study of browsing. Traditionally browsing has attracted little research interest, perhaps because it is nebulous and unquantifiable. This paper presents a way to change that, in the form of a transaction log analysis of browsing using circulation data from a large academic library consortium; an entirely new but reproducible technique.

A pair-wise comparison of books by both location and loan date demonstrates a clear neighbour effect in borrowing, a finding completely novel in the literature. This effect is independent of

both book topic subject and time. This neighbour effect strongly suggests browsing, and is in line with both the information seeking literature, and with information seekers' self-reports about browsing.

Further examination of the data shows that the browsing effect is significant—a nearby book being loaned increases the probability of a loan by at least 9-fold. This means that browsing is occurring regularly and that users are effective even in the limited and limiting context of physical shelves with arcane classification systems. It also gives us a benchmark for testing other browsing systems: less than a 9-fold increase in a pairwise comparison matched by location means a poorer return than a physical library.

With this clear confirmation of the value of browsing for books, and in the context of increasing ebook usage, providing good opportunities for online browsing is more important now than ever. To fully understand the browsing process and users browsing needs, though, requires considerable future work.

### 6.2 Future Work

This study, like all transaction log analyses can only identify user behaviour that generates a system event, in this case a loan. There are many steps between searching or browsing and making that loan, however, that are opaque to this study: examining the shelves, close examination of individual books, even how books are encountered.

Similarly we do not know how satisfactory the browsing experience was for the information seekers who created these loans, nor do we understand what features of the shelves are imperative and which could be improved.

Close observational study in both online book selection environments and physical spaces would go some way toward clarifying what information seekers want from book browsing, and finally informing the design of truly good, usable online browsing systems.

## 7. ACKNOWLEDGEMENTS

This paper contains information from OhioLINK Circulation Data (<http://www.oclc.org/research/activities/ohiolink/circulation.htm>) which is made available by OCLC Online Computer Library Center, Inc. and OhioLINK under the ODC Attribution License (<http://www.oclc.org/research/activities/ohiolink/odcby.htm>). Thanks to Dr. Ed O'Neill from the OCLC for answering queries and providing further background to the data

Finally, thanks to George Buchanan for advice on statistics, and to the reviewers for their insightful comments and recommendations.

## 8. REFERENCES

- [1] Baeza-Yates, R., Hurtado, C. and Mendoza, M. Query Recommendation Using Query Logs in Search Engines. In *Proc. EDBT 2004 Workshops*. Springer Berlin Heidelberg, 2005. 588-596.
- [2] Bates, M. J. The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, 13, 5, 407-424 1993.
- [3] Bates, M. J. What is browsing--really? A model drawing from behavioural science research. *Information Research*, 12 2007.
- [4] Berg, S. A., Hoffmann, K. and Dawson, D. Not on the Same Page: Undergraduates' Information Retrieval in Electronic and Print Books. *The J Acad Libr*, 36, 6, 518-525 2010.

- [5] Blandford, A., Rimmer, J. and Warwick, C. Experiences of the Library in the Digital Age. In *Proc. CCDT 06* Tavros, Greece, 23-25 November. Foundation of the Hellenic World, 2006.
- [6] Borgman, C. L. Why are online catalogs *still* hard to use? *JASIST*, 47, 7, 493-503 1996.
- [7] Borgman, C. L., Hirsh, S. G., Walter, V. A. and Gallagher, A. L. Children's searching behavior on browsing and keyword online catalogs: The Science Library Catalog project. *JASIST*, 46, 9, 663-684 1995.
- [8] Bowker, G. C. and Star, S. L. *Sorting things out: classification and its consequences*. MIT Press, Boston, MA USA, 2000.
- [9] Bryant, J., Matthews, G. and Walton, G. Academic libraries and social and learning space. *Journal of Librarianship and Information Science*, 41, 1, 7-18 2009.
- [10] Buchanan, G., Cunningham, S. J., Blandford, A., Rimmer, J. and Warwick, C. Information Seeking by Humanities Scholars. In *Proc. ECDL 05* Vienna, Austria, September 18-23. Springer, 2005. 218-229.
- [11] Buchanan, G. and McKay, D. In the Bookshop: Examining Popular Search Strategies. In *Proc. JCDL 11* Ottawa, Canada, June 14-17. ACM, 2011.
- [12] Burrows, T. A machine for browsing: beyond the "single search box". *OCLC Systems & Services*, 28, 1, 56-58 2012.
- [13] Casey, M. and Stephens, M. It's Fine to Drop Dewey. *Library Journal*, 134, 12, 19 2009.
- [14] Christianson, M. and Aucoin, M. Electronic or print books: Which are used? *Library Collections, Acquisitions, and Technical Services*, 29, 1, 71-81 2005.
- [15] Cooksey, E. B. Too Important to Be Left to Chance—Serendipity and the Digital Library. *Science & Technology Libraries*, 25, 1-2, 23-32 2004.
- [16] Crabtree, A., Twidale, M. B., O'Brien, J. and Nichols, D. M. Talking in the Library: Implications for the Design of Digital Libraries. In *Proc. DL 97* Philadelphia, PA, USA. ACM Press, 1997. 221-228.
- [17] Craswell, N. and Szummer, M. Random walks on the click graph. In *Proc. SIGIR 07* Amsterdam, The Netherlands. ACM, 2007. 239-246.
- [18] Cunningham, S., Vanderschantz, N., Timpany, C., Hinze, A. and Buchanan, G. Social Information Behaviour in Bookshops: Implications for Digital Libraries. In *Proc. TPDL '13*, Valetta, Malta. Springer Berlin Heidelberg, 2013. 84-95.
- [19] Cunningham, S. J. How children find books for leisure reading: implications for the digital library. In *Proc. JCDL '11* Ottawa, Ontario, Canada. ACM, 2011. 431-432.
- [20] Cunningham, S. J., Knowles, C. and Reeves, N. An ethnographic study of technical support workers: why we didn't build a tech support digital library. In *JCDL '01* Roanoke, Virginia, United States, June 24-28. ACM, 2001.
- [21] Cunningham, S. J., Reeves, N. and Britland, M. An ethnographic study of music information seeking: implications for the design of a music digital library. In *Proc JCDL 03* Houston, Texas. IEEE, 2003. 5-16.
- [22] Fast, K. V. and Campbell, D. G. "I still like Google" University student perceptions of searching OPACs and the web. *ASIST Proceedings*, 41, 1, 138-146 2004.
- [23] Foster, A. and Ford, N. Serendipity and information seeking: an empirical study. *J Doc*, 59, 3, 321-340 2003.
- [24] Gerwe, P. and Viles, C. L. User effort in query construction and interface selection. In *Proc. DL 00* San Antonio, Texas, USA, June 2-7. ACM, 2000.
- [25] Gorman, M. The Longer the Number the Smaller the Spine. *American Libraries*, 12, 8, 498 1981.
- [26] Gross, J. and Sheridan, L. Web scale discovery: the user experience. *New Library World*, 112, 5, 236-247 2011.
- [27] Hamblen Amazon: E-books now outsell print books. *Computerworld*, June 14 2011.
- [28] Hancock-Beaulieu, M. A comparative transaction log analysis of browsing and search formulation in online catalogues. *Program: electronic library and information systems*, 27, 3, 269-280 1993.
- [29] Hancock-Beaulieu, M. Evaluating the impact of an online library catalogue on subject searching at the catalogue and at the shelves. *J Doc*, 46, 4, 318-338 1993.
- [30] Herlocker, J. L., Konstan, J. A., Terveen, L. G. and Riedl, J. T. Evaluating collaborative filtering recommender systems. *ACM TOIS*, 22, 1, 5-53 2004.
- [31] Hinze, A., McKay, D., Vanderschantz, N., Timpany, C. and Cunningham, S. J. Book selection behavior in the physical library: implications for ebook collections. In *Proc. JCDL '12* Washington, DC, USA. ACM, 2012. 305-314.
- [32] Huvila, I. "Library users come to a library to find books": The structuration of the library as a soft information system. *J Doc*, 69, 5, 715-735 2013.
- [33] Jiao, Q. G. and Onwuegbuzie, A. J. Reading ability as a predictor of library anxiety. *Library Review*, 52, 4, 159-169 2003.
- [34] Jones, S., Cunningham, S. J. and McNab, R. An Analysis of Usage of a Digital Library. In *Proc. ECDL 98* Heraklion, Crete, September 21-23. Springer, 1998. 261-277.
- [35] Kamvar, M. and Baluja, S. A large scale study of wireless search behavior: Google mobile search. In *Proc. CHI 06* Montreal, Quebec, Canada. ACM, 2006. 701-709.
- [36] Kleiner, E., Rädle, R. and Reiterer, H. Blended shelf: reality-based presentation and exploration of library collections. In *Proc. CHI* Paris, France, April 27-May 2. ACM, 2013.
- [37] Kuhlthau, C. C. Inside the Search Process: Information Seeking from the User's Perspective. *JASIST*, 42, 5, 361-371 1999.
- [38] Kules, B., Capra, R., Banta, M. and Sierra, T. What do exploratory searchers look at in a faceted search interface? In *Proc. JCDL 09* Austin, TX, USA. ACM, 2009. 313-322.
- [39] Lau, E. P. and Goh, D. H.-L. In search of query patterns: A case study of a university OPAC. *Information Processing & Management*, 42, 5, 1316-1329 2006.
- [40] 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, Oakland, CA, USA, 2011.

- [41] Liesaputra, V. and Witten, I. H. Seeking information in realistic books: a user study. In *Proc. JCDL '08 Pittsburgh PA, PA, USA*. ACM, 2008. 29-38.
- [42] Loizides, F. and Buchanan, G. Towards a Framework for Human (Manual) Information Retrieval. In *Proc. IRFC 13* Springer Berlin Heidelberg, 2013. 87-98.
- [43] Losee, R. M. The relative shelf location of circulated books: A study of classification, users, and browsing. *LRTS*, 37, 2, 197-209 1993.
- [44] Mahoui, M. and Cunningham, S. J. Search Behavior in a Research Oriented Digital Library. In *Proc. ECDL 01* Darmstadt, Germany, September 4-9. Springer, 2001. 13-24.
- [45] Makri, S. and Blandford, A. Coming across information serendipitously: Part 1 – A process model. *J Doc*, 68, 5, 5-5 2012.
- [46] Makri, S., Blandford, A., Gow, J., Rimmer, J., Warwick, C. and Buchanan, G. A library or just another information resource? A case study of users' mental models of traditional and digital libraries. *JASIST*, 58, 3, 433-445 2007.
- [47] Makri, S., Blandford, A., Woods, M., Sharples, S. and Maxwell, D. "Making my own luck": Serendipity strategies and how to support them in digital information environments. *JASIST*, to appear 2014.
- [48] Malama, C., Landoni, M. and Wilson, R. Fiction Electronic Books: A Usability Study. In *Proc. ECDL '04*. Springer Berlin / Heidelberg, 2004. 69-79.
- [49] Marchionini, G. *Information Seeking in Electronic Environments*. Cambridge University Press, UK, 1995.
- [50] Marshall, C. C. *Reading and Writing the Electronic Book*. Morgan & Claypool, Chapel Hill, NC USA, 2010.
- [51] Marshall, C. C. and Bly, S. Turning the page on navigation. In *Proc. JCDL'05 Denver, CO, USA*. ACM, 2005. 225-234.
- [52] McKay, D. Gotta keep 'em separated: Why the single search box may not be right for libraries. In *Proc. CHINZ 11* Hamilton, New Zealand, 4-5 July. ACM, 2011. 109-112.
- [53] McKay, D. and Conyers, B. Where the streets have no name: how library users get lost in the stacks. In *Proc CHINZ 10* Auckland, New Zealand, 8 July. ACM, 2010. in press.
- [54] McKay, D., Hinze, A., Heese, R., Vanderschantz, N., Timpany, C. and Cunningham, S. J. An Exploration of ebook Selection Behavior in Academic Library Collections. In *Proc TPDFL '12 Paphos, Cyprus*, 23-27 September. Springer, 2012. 13-24.
- [55] McKay, D., Shukla, P., Hunt, R. and Cunningham, S. J. Enhanced browsing in digital libraries: three new approaches to browsing in Greenstone. *JoDL* 4, 4, 283-297 2004.
- [56] McKenzie, P. J. A model of information practices in accounts of everyday-life information seeking. *J Doc*, 59, 1, 19-40 2003.
- [57] Meadow, K. and Meadow, J. Search Query Quality and Web-Scale Discovery: A Qualitative and Quantitative Analysis. *College & Undergraduate Libraries*, 19, 2-4, 163-175 2012.
- [58] Moore, P. Information Problem Solving: A Wider View of Library Skills. *Contemp Educ Psych*, 20, 1, 1-31 1995.
- [59] Nicholas, D., Huntington, P., Jamali, H. R. and Tenopir, C. 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 2006.
- [60] Nordlie, R. "User Revelation"—A Comparison of Initial Queries and Ensuing Question Development in Online Searching and in Human Reference Interactions. In *Proc. SIGIR 99 Berkeley, CA, USA*, August 15-19. ACM Press, 1999. 11-18.
- [61] O'Neill, E. T. and Gammon, J. A. Consortial Book Circulation Patterns: The OCLC-OhioLINK Study. *College & Research Libraries* In Press.
- [62] Ooi, K. *How Adult Fiction Readers Select Fiction Books in Public Libraries: A Study of Information Seeking in Context*. Masters, Victoria University of Wellington, Wellington, New Zealand, 2008.
- [63] Pearce, J., Chang, S., Alzougool, B., Kennedy, G., Ainley, M. and Rodrigues, S. Search or explore: do you know what you're looking for? In *Proc. OZCHI 11*, Canberra, Australia. ACM, 2011. 253-256.
- [64] Reutzell, D. R. and Gali, K. The Art of Children's Book Selection: A Labyrinth Unexplored. *Reading Psychology*, 19, 1, 3-50 1998.
- [65] Rowlands, I. and Nicholas, D. Understanding Information Behaviour: How Do Students and Faculty Find Books? *J Acad Libr*, 34, 1, 3-15 2008.
- [66] Rowlands, I., Nicholas, D., Jamali, H. R. and Huntington, P. What do faculty and students really think about e-books? *Aslib Proceedings*, 59, 6, 489-511 2007.
- [67] Saarinen, K. and Vakkari, P. A sign of a good book: readers' methods of accessing fiction in the public library. *J Doc*, 69, 5, 736-754 2013.
- [68] Saarti, J. Feeding with the spoon, or the effects of shelf. *Information Services & Use*, 17, 2/3, 159 1997.
- [69] Sadeh, T. User Experience in the Library. *New Library World*, 109, 1/2, 7-24 2007.
- [70] Spink, A., Wolfram, D., Jansen, M. B. J. and Saracevic, T. Searching the web: The public and their queries. *JASIST*, 52, 3, 226-234 2001.
- [71] Stelmaszewska, H. and Blandford, A. From physical to digital: a case study of computer scientists' behaviour in physical libraries. *JoDL*, 4, 2, 82-92 2004.
- [72] Svenonius, E. *The Intellectual Foundation of Information Organization*. MIT Press, Boston, MA, USA, 2000.
- [73] Talja, S. and Maula, H. Reasons for the use and non-use of electronic journals and databases: A domain analytic study in four scholarly disciplines. *J Doc*, 59, 6, 673-691 2003.
- [74] Thudt, A., Hinrichs, U. and Carpendale, S. The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In *Proc. CHI 11 Austin, Texas, USA*. ACM, 2012. 1461-1470.
- [75] Wallace, P. M. How do patrons search the online catalog when no one's looking? Transaction log analysis and implications for bibliographic instruction and system design. *Reference & User Services Quarterly*, 33, 2, 239-253 1993.
- [76] Zhang, Y. and Moffat, A. Some observations on user search behavior. *Australian Journal of Intelligent Information Processing Systems*, 9, 2, 1-8 2006.