The things we talk about when we talk about browsing: an empirical typology of library browsing behaviour

Dana McKay, University of Melbourne, Parkville VIC 3010, Australia. Dana.mckay@unimelb.edu.au, phone: +61420422215, fax +61393494596 (corresponding author)

Shanton Chang, University of Melbourne, Parkville, VIC 3010, Australia. Shanton.chang@unimelb.edu.au, phone: +61383441583, fax +61393494596

Wally Smith, University of Melbourne, Parkville, VIC 3010, Australia, wsmith@unimelb.edu.au, phone: +61383441494, fax: +61393494596

George Buchanan University of Melbourne, Parkville, VIC 3010, Australia. George.buchanan@unimelb.edu.au, phone: +61383443775, fax +61393494596
Abstract
Libraries increasingly offer much of their collection online, rendering it invisible or unavailable to readers who, for reasons of information experience, prefer to browse the shelves. While the evidence that shelf browsing is an important part of information behaviour is increasing, information browsing as a behaviour is somewhat of a black box (in contrast to web browsing, which is relatively well understood). It seems likely from early work that browsing is not, in fact, a monolithic behaviour, but rather a set of behaviours and goals. The typologies presented in these works, however, are too high level to offer much insight into what support is needed for successful online browsing. In contrast, a recent spate of speculative browsing technologies meet some browsing needs, but offer little theoretical understanding of how systems that support browsing. The major contribution of this paper is a new typology of library browsing behaviour based on recent observations of browsing behaviour in libraries. The secondary contribution is an understanding of the interface features that would support these types of information browsers in an online environment.

Introduction
Technology is changing the way libraries deliver information. In Australia and New Zealand in 2014, ebooks represented over 90% of all books held in academic libraries; in the UK ebooks went from 10% of holdings in 2010 to 20% of holdings in 2014. Despite library users’ mixed feelings about ebooks, they are an increasing proportion of what libraries offer. Readers raise two major reservations about using ebooks: readability and browsability (Makri et al., 2007; Marshall, 2010; Pearson, Buchanan, & Thimbleby, 2010; Rowlands, Nicholas, Jamali, & Huntington, 2007). The challenges of readability are manifold, including poor support for annotation, limited navigability, and the frustrations caused by digital rights management software (Marshall & Bly, 2005; Pearson et al., 2010; Tonkin, 2010); however the focus of the work presented here is browsing.

Browsing is a clear advantage of visiting a library in person; study after study has quoted users commenting on the value of the library shelves as a discovery resource (Blandford, Rimmer, & Warwick, 2006; Makri et al., 2007; Stelmaszewska & Blandford, 2004). While shelf browsing might seem anachronistic in the age of the internet, a survey in 2013 demonstrated that over half of the users of a German library considered it an important component of their information seeking process (Kleiner, Rädle, & Reiterer, 2013), a finding supported by our own recent work (McKay, Smith, & Chang, 2014; McKay, Smith, & Chang, 2015). The ebook revolution, in its present form, represents a near-complete return to closed-stack libraries, with materials only accessible via the medium of search; browsing is simply not available.

The loss of the option to browse means those seeking ebooks must rely on search, which is notoriously poor for supporting imprecisely defined information needs (Borgman, 1996; Kuhlthau, 1991; Marchionini, 1997, Belkin, Oddy, & Brooks, 1982) and supports serendipity poorly (Foster & Ford, 2003). Given the importance of serendipity and browsing to information work, they are information seeking strategies we lose at our peril (Cooksey, 2004; Foster & Ford, 2003; Makri & Blandford, 2012a, 2012b).

---


2 https://www.sconul.ac.uk/sites/default/files/documents/Analysis%20_Loans%20ebooks%20visits%20June%202015.pdf
In response to this peril, recent times have seen a rash of online browsing technologies being tested either in research settings (e.g. Kleiner et al., 2013; Pearce & Chang, 2014; Thudt, Hinrichs, & Carpendale, 2012), or being released by commercial organizations (e.g. the ExLibris virtual shelf\(^3\)). None of these interfaces has (yet) seen widespread uptake or approbation; there is simply no tool for browsing that has even a modicum of the recognition that Google has for search. Given that browsing is a poorly understood behaviour (Joranson, VanTuyl, & Clements, 2014; Tonkin, 2010), it seems likely that a big part of the reason these tools are not seeing widespread use is that they fundamentally do not facilitate what users would recognise as browsing in a meaningful way—a hypothesis for which there is some evidence (McKay, Buchanan, & Chang, 2018).

Browsing has been treated as something of a black box by the major information seeking models—they acknowledge it happens, but they do not address what it is (Ellis, 1989; Kuhlthau, 1991; Marchionini, 1997). Early literature suggests that browsing might be more than one behaviour: a number of typologies of browsing behaviour exist (Apted & Choo, 1971; Celoria, 1969; Herner, 1970; Rice, McCreadie, & Chang, 2001). Each of these typologies comes from a time before the internet (Rice et al.’s data is from 1991), and only Rice et al.’s is empirically grounded.

The work presented in this paper aims to address some of the gaps in our understanding of browsing by developing an internet-era, empirically grounded typology of browsing based on observations of library shelf browsing. We then use this typology to make design suggestions for online browsing, a goal that has become pressing with the ebook revolution. The rest of the paper is divided up as follows: first we describe the literature relating to browsing; next we describe the method we used to derive the browsing taxonomy presented here. Third we describe the taxonomy, and its implications for future browsing interfaces. Finally we describe our taxonomy and its implications in the context of the literature, and draw conclusions.

**Background**

In this section we provide the literature background to our work. First we position a definition of browsing in the major information models, and examine why it is important. Next, we examine the empirical literature on library shelf use, addressing how our study extends the literature. Finally, we examine existing typologies of browsing noting specifically the theoretical contribution of the work presented here.

**On Browsing and Its Importance**

Browsing is a fundamental component of information behaviour, appearing in every major model of information seeking. Marchionini defines browsing in contrast with search, noting that the information environment is a key component of the behaviour, and that it is informal and opportunistic (Marchionini, 1997). Kuhlthau’s ‘exploration’ (Kuhlthau, 1991), the closest analogue in her model, notes that information seekers need to explore and that they may find information systems unhelpful, but does not explain what this exploration looks like in practice. Ellis talks about browsing as ‘undirected’ information seeking and notes that it should be supported online (Ellis, 1989). His approach to supporting it online is relatively simplistic, saying that any information held about a collection should be browsable, but he does not describe the interface this should entail. Wilson’s model (Wilson, 1999) takes a higher level overview of the information seeking process than the other three and does not address browsing specifically, but notes that there can be a high

---

\(^3\) https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Technical_Guide/010The_PNX_Record/150The_Browse_Section/020Virtual_Shelf_Tab
degree of the ambiguity in information needs—ambiguity that the other models suggest is best resolved by browsing.

The definitions of browsing given above are somewhat contested, but perhaps the most comprehensive definition to date is the one Bates gave in her overview of literature from a range of disciplines (Bates, 2007). It is given as a series of steps:

‘1) glimpse a scene, 2) home in on an element of a scene visually and/or physically (if two or more elements are of interest, they are examined serially, not in parallel), 3) examine item(s) of interest, 4) physically or conceptually acquire or abandon examined item(s)’.

Bates makes it clear elsewhere in this paper that the scene must be ‘large and interesting’. This definition is an excellent working description of browsing in situations where information decisions must be based on what Marchionini termed surrogate representations of the information (Marchionini, 1997); situations where the items being browsed cannot be ingested in their entirety.

Examples of surrogate browsing include video browsing and—pertinent to the work presented here—browsing the library shelves. Bates’ definition does, however, exclude much of what we refer to as browsing in the information seeking literature: Web browsing, for example, is information acquisition rather than seeking, and much more like Pirolli’s information foraging (Pirolli, 1997) where the item viewed is an information resource in itself. Not all online browsing will be like this, though—ebook browsing, music browsing, and video browsing will all require surrogates online just as they do in the physical realm.

In a behavioural attempt to define browsing, both Marchionini and Kwasnik have described some of the activities involved in browsing—Kwasnik notes six activities (Kwasnik, 1992), and Marchionini four (Marchionini, 1997). More recently the authors of this paper have identified a number of browsing behaviours (McKay, Chang, & Smith, 2017). While supporting these behaviours and activities might well improve online browsing support generally, understanding the goals and needs information seekers are addressing while they engage in these activities is likely to provide further insight into what an effective system might look like. All the major models note that browsing is under-studied and poorly supported online (Ellis, 1989; Joranson et al., 2014; Kuhlthau, 1991; Marchionini, 1997). This dearth of study is particularly noticeable when we compare browsing with search, which has been the subject of continued and focused study since the Cranfield experiments in the 1950s (Baeza-Yates & Ribeiro-Neto, 1999).

The shelves have been described as a near-perfect browsing engine (Kleiner et al., 2013) (based on Bates’ definition above); until the ebook revolution the imperative to support online browsing has been low. Now, ebook users must access books using search, a strategy poorly suited to loosely defined information needs. In the past these searchers may have consulted reference librarians to help them turn a need into a precisely worded successful query (Crabtree, Twidale, O'Brien, & Nichols, 1997; Nordlie, 1999). A modern information seeker sitting at home at midnight, however, is unlikely to have access to such a resource. The loss of browsing for such a large part of any library’s collection results in a loss of serendipitous discovery (Foster & Ford, 2003; Makri et al., 2007) and higher cognitive burdens for information seekers (Ellis, 1989; Marchionini, 1997). In a time where more information than ever before is online, the need to understand and support online surrogate browsing has never been more pressing.

In this paper we attempt to address both the poor understanding of browsing—and the poor online support for it—by observing readers browsing in a modern library setting.
Library shelf use

Many previous studies have, intentionally or incidentally, investigated information seeking using library shelves. A spate of observational studies of library users in the mid-2000s demonstrated that library shelves are a key component of book seeking (Blandford et al., 2006; Makri et al., 2007; Stelmaszewskas & Blandford, 2004). The key finding from these studies in terms of the work presented here is that readers value shelf browsing highly—some to the point that they eschew ebooks (Hinze, McKay, Vanderschantz, Timpany, & Cunningham, 2012). Readers rely on shelves for serendipity, and to expand on the offerings of library catalogues. Readers also use information available at the shelves—such as book size, dust (indicating infrequent use) and visual cues from book covers to assist in their book selection processes (Ooi, 2008; Saarinen & Vakkari, 2013). These studies identified problems with the shelf experience, too, though—books could be missing, or checked out, and as noted in previous research, semantically related books can be shelved at geographically distant locations (Losee, 1993).

Rice et al examined data from a 1991 study of library use, and came up with a descriptive framework of browsing behaviour that they presented in their book. This behaviour, though, is based on pre-internet library use and varies in the detail with which it describes behaviour (Rice et al., 2001). Reutzel and Gali examined children’s library shelf browsing in 1997 and came up with a framework of behaviours (Reutzel & Gali, 1998). These behaviours were low level physical behaviours, however, and did not describe the children’s information seeking goals or strategies. Hinze applied this framework to adult shelf use in 2012 (Hinze et al., 2012), describing the shelf interactions present in accessing a single book. This work was refined in 2017, when the authors of this paper examined complete browsing interactions to generate a new list of actions (McKay et al., 2017).

Many previous studies have confirmed the influence that shelf layout has on information seeking. Beaulieu’s 1997 study of readers choosing books noted that shelf interactions prompted readers to select additional books over half the time (Hancock-Beaulieu, 1993). This effect may well be amenable to influence; Saarti’s work demonstrated that rearranging the shelves can affect how much browsing actually takes place (Saarti, 1997). An early statistical analysis of circulation of small parts of a library’s collection confirmed that shelf location affected not just how much browsing happened, but what was borrowed (Losee, 1993).

While shelf browsing may seem like an activity from a quaint and nostalgic past, Kleiner’s work has shown that it is valued by approximately half of all academic library users (Kleiner et al., 2013). The statistical effects on borrowing demonstrated by Losee were replicated by the authors of this paper over whole library collections in 2014. Further work demonstrated that ebooks do not show this evidence of browsing (McKay et al., 2014; McKay, Smith, et al., 2015). Our work follows in a tradition of using physical to generate digital interface designs (e.g. Pearson, Buchanan, Thimbleby, & Jones, 2012, Cunningham, Reeves, & Britland, 2003). As with this previous work understanding the physical behaviour—shelf browsing—is necessary because directly replicating the physical realm is impossible without virtual reality equipment, and ineffective where a virtual reality approach has been trialled (Almeida, Cubaud, Dupire, Natkin, & Topol, 2006).

While previous work has demonstrated that shelf browsing is an important part of library information seeking, what readers actually do, what their motives and intentions are, and what a browsing session looks like has not yet been investigated. This paper aims to address that gap, providing a taxonomy of browsing behaviours.
Other Information Seeking Typologies

One of the fundamental advances in understanding search behaviour—and how to support it online—came with the advent of Broder’s search taxonomy (Broder, 2002). This taxonomy is empirically grounded, and offers three types of search, each of which needs different behaviour from a system. Our understanding of browsing is considerably more limited: all the existing typologies are pre-internet, and only one has any clear empirical basis.

The earliest browsing typology is from 1969; in this Celoria distinguishes professional browsing—repeated examination of the shelves and deliberate serendipity seeking—from recreational browsing, which is merely wandering in the stacks (Celoria, 1969). In 1970 Herner divided browsing into three classes: directed, undirected and semidirected browsing (Herner, 1970). Undirected browsing is the same as Celoria’s recreational browsing, and semidirected browsing represents deliberate serendipity seeking. Directed browsing has a clear goal, suggesting it might be excluded from, for example, Marchionini’s description of browsing. In 1971, Apted divided browsing into two categories—specific browsing, and general purposive browsing (Apted & Choo, 1971). The difference between these two types is the end goal: specific browsing is aimed at locating a clear target, general purposive browsing is wandering in the hope of finding non-specific new information. Marchionini divides browsing based on what is being browsed, referring to across- and within-document browsing. When examining library browsing, this distinction is not useful—book selection usually involves both (Marchionini, 1997). Finally Rice et al. derive four browsing types from a range of literature—looking for a specific item; looking for items that can be grouped by characteristic in some way; collection evaluation or understanding; and incidental browsing (Rice et al., 2001). As with Marchionini’s division, a typical visit to the library may include multiple browsing types from Rice’s model.

Most of these models are a continuum, with readers seeking a specific item at one end, and engaging in very loosely directed information seeking at the other; what is in between varies between typologies, as does the role of serendipity. The majority of these models predate the internet, and there is not a clear empirical underpinning to any of them. In this paper we attempt to generate a finer-grained, empirically grounded typology that addresses the role of serendipity. We further use this typology to point to avenues for developing better systems for online browsing.

Method

There is a long tradition of examining information behaviour in the physical realm to inform the design of digital support, for example (Cunningham et al., 2003; Cunningham, Rogers, & Kim, 2017; Pearson et al., 2012; Stelmaszewska & Blandford, 2004); this work engages in that tradition. The results described in this paper are derived from two studies. The first was a series of observations of shelf use, coupled with interviews. The first analysis of this data, presented in (McKay et al., 2017), described the individual actions readers took while browsing. This paper presents a new analysis, examining the strategies and goals of readers in relation to their browsing.

The second study—which is presented for the first time in this paper—was designed to triangulate results from the first and consisted of interviews taken at the point where readers were borrowing books. Both studies were approved by the Human Research Ethics Committee at the University of Melbourne. In this section we first give the details of each study, then describe our approach to analysing the data for the purposes of this paper.
Observational study
This study took place in two locations: the library of a small, research-active university in Melbourne, Australia, and a large, mandated-deposit reference library in the same city (the State Library of Victoria). The selection of the two sites was to avoid the impact that individual building architecture or user populations may have had on our results; given that results are drawn from two sites we believe they are likely to generalise to other academic and reference libraries.

This study resulted in 31 observations, 15 at the university site and 16 at the State Library. Participants were evenly divided along gender lines, and 23 were students (14 from the university site and 7 from the State Library). Observations continued until saturation was reached, i.e. no new behaviours or intentions were recorded.

Our approach follows a long tradition of observational studies of the library shelves (Blandford & Furniss, 2006; Hancock-Beaulieu, 1993; Hinze, Alqurashi, Vanderschantz, Timpany, & Alzahrani, 2014; Hinze et al., 2012; Makri et al., 2007; Stelmaszewska & Blandford, 2004). Each of these studies followed a similar protocol: observe readers, and ask them about their information practices and their experience of the library as an information system. Our study took the same approach. The difference was, our focus was specifically browsing behaviour, and we followed readers from the moment they consented to our observation (as they approached the shelves) until they completed their information seeking process. The protocol for this study was to observe participants as they examined the shelves, noting their physical behaviours and interactions with books. Participants were questioned about some aspects of their behaviour to ensure we understood their goals and intentions. Participants were further questioned about what their information goals had been when they came to the library; whether they had used the library catalogue; what their planned information seeking strategy had been; and whether they had found what they were looking for. Once the observation had finished, the observer assisted any participants who had not found desired specific materials in either locating those materials or confirming their absence from the library (though participants were not told ahead of time that this would happen). Only a single participant declined to participate in this study.

Observations and interviews were recorded using handwritten notes, and tallies of books and shelves investigated and interacted with. Immediately upon completion of each observation, data was transcribed into an excel spreadsheet, noting interactions and writing a ‘browsing story’ for each participant, akin to Makri’s serendipity stories (Makri, Blandford, Woods, Sharples, & Maxwell, 2014).

This study was analysed as described below before the commencement of the second study.

Interview study
We know that participants’ actions and their interpretations of their actions can be at odds (Rogers, Sharp, & Preece, 2007). The second study reported in this paper was designed to both understand participants’ own interpretations of their shelf browsing experiences and to triangulate findings from the first study against browsers’ own accounts of their experiences.

In this study, participants were approached as they borrowed books from the university library described above. Book borrowing represents a certain amount of investment in using the books being borrowed; users prefer to minimize borrowing to avoid fines and carrying home heavy materials (Hinze et al., 2012). In this study it was used not just as a metric of interest, but also as a sign that the information seeking process in which these readers were engaged was completed, at
least in the short term. Because this protocol relied on borrowing, it could not be replicated in the State Library, which does not allow borrowing.

The interviewing researcher in this study waited near the self-checkout machines in the library (the only way to borrow) for readers borrowing more than one book. This strategy was designed to maximise the probability that these information seekers would have done at least some browsing, rather than approaching the shelves to borrow only books they had searched for and then leaving. In contrast to the first study, four prospective participants declined to participate in this study; while no reason was asked for they all commented that they were leaving the library and rushing to move on to their next task. Interviews were conducted on weekdays during term time during business hours. While this slightly reduced the chance of interviewing participants who have been engaged in browsing, which is more common at weekends, (McKay, Buchanan, & Chang, 2015), it is the busiest time for the library and maximized recruitment opportunities.

Once participants had agreed to be interviewed, the call numbers of the books they had borrowed were documented and they were interviewed about their experience of selecting those books. Interview questions established whether participants had searched the library catalogue; how many of the books they were borrowing had been identified during search; how many shelves they remembered examining; whether they had examined any books that they had not selected and (in their own words) what their book selection process had been. Interviews typically lasted between 5 and 8 minutes, and were audio recorded.

As in the previous study, interviews continued until the interviewing researcher was satisfied that saturation had been reached, a total of 13 interviews. While this is a relatively small number, it is not incommensurate with other studies of this type—Makri et al. (2007) had 8 participants, for example, and Stelmaszewsk & Blandford (2004) observed 9 and interviewed a further 5. Given that this study was being used to test the results of the previous study, saturation was reached rapidly.

After interviewing participants, interviews were transcribed into an excel spreadsheet noting the presence or absence of search; how many shelves readers believed they had examined; how many books they had examined and borrowed; and how many of the books borrowed had been searched for. As with the previous study, a browsing story was written for each participant, narrating their experience as much as possible in their own words. This spreadsheet formed the basis for data analysis.

**Data analysis**

As mentioned above, data for the observational study was analysed first. This data was analysed by two researchers using an inductive approach similar to grounded theory (Glaser & Strauss, 1967), however it was analysed post-collection due to the availability of the second researcher. In the first instance, each researcher independently examined the data, doing inductive coding on the behaviours and the overarching strategies used by browsers. Both researchers have a strong background in information seeking behaviour, so classifications were influenced by their knowledge of the literature.

After an initial pass, coding schemes were compared and refined, and the data was re-coded using the resulting classification scheme. This classification scheme was discussed and refined, then applied to the data. A further refinement and application of the scheme resulted in an inter-rater agreement of over 90% over the entire classification, and 90.3% for the six browsing types. The same two researchers used the browsing stories generated by the interviews to classify these participants by browsing type, though both researchers were aware they may see a new type during this
classification. No new types were identified, and in classification the researchers reached a 100% agreement for which type each interview represented.

Finally, the researchers discussed extensively the defining features of each browsing type, generating a classification scheme of those types on four axes (serendipity; presence or absence of a sought item; presence or absence of search; and how tightly defined the information goal was). The researchers independently classified the browsing types according to this scheme; all classes were found to be clear and easy to apply. Agreement of 100% was reached.

Findings: A Typology of Browsing

In this section we first describe the six types of shelf browsing identified in our study, along with their prevalence in our data, and their implications for shelf browsing. We then describe their relationship to one another along four axes. In all descriptions of individual readers, information seeking topics have been altered subtly to protect reader privacy.

The types

Our research identified 6 types of shelf-based information seeking behaviours, 5 of which involved some measure of browsing. In contrast to the behaviours described in (McKay et al., 2017), which describes the individual actions that make up a browsing session, the six behaviours presented here represent the overarching goals and strategies employed by the information seekers.

The six behaviours are listed below, in order of how well they might be supported by traditional interactive search, from most to least supported. It should be emphasized that these types represent episodes of behaviour, rather than types of individual: while individuals may have preferred information seeking styles among the types, the types are influenced by information needs and environments (as we would expect), and so are not simply preference driven. The names for the behaviours are derived from digital library research; the aim of this work, after all, is to improve browsing in the digital library.

Grab-and-go

Grab-and-go behaviour is the selection of a book or books pre-identified by searching that actively excludes the consideration of any alternative books. This behaviour is not browsing, it is acquisition; all other behaviours in the typology do involve some browsing.

**Motivation:** This process is focused on selecting a specific item or items, much like the bookshop readers observed in (Buchanan & McKay, 2011), from which the term ‘grab-and-go’ is drawn. If the item or items these readers were looking for was missing from the shelves, they would leave empty handed rather than selecting an alternative. This is akin to a classic known-item search in (Numminen & Vakkari, 2009), and to known-item finding in (Björneborn, 2008) These users accounted for 22% of all participants in this study—10 out of 45 overall, but 9 of 32 in the observational part of the study.

**Behaviours:** Readers engaged in this process will first search the catalogue, then proceed as directly to the relevant shelves as their knowledge of their library allows, then select their books before moving on to the next task. In this process the shelves and the distraction they imply are an impediment to information seeking.

**Implications for online browsing:** Assuming no reason for avoiding the use of ebooks, this process is perfectly supported by relevance ranked search; much better than it is supported by any process involving the library shelves.
Satisficing
Satisficing is the grudging selection of a semantically and topically similar book or books, because pre-identified target books were not available.

**Motivations:** Satisficing, first described in (Agosto, 2002) as selecting an imperfect information resource rather than investing further time to find the perfect resource, is born of similar motivations to grab-and-go browsing. Readers who are satisficing want to access a small number of pre-identified items. Where satisficing varies from grab-and-go browsing is that satisficing readers will accept a near-neighbour substitute for books they cannot locate on the shelves. The classic example of this from our data is the reader who—when she was unable to locate a biology textbook prior to her tutorial—selected a different textbook called ‘Biology’ saying “I just want to pass”. We observed 6 of these readers and interviewed a further 3, for a total of 20% of all participants. While this behaviour appears similar to browsing for a topically similar item, it is more complex. Topic certainly matters, but there are often other considerations; in the example of the woman seeking a textbook, only another textbook would do, rather than any book on biology.

**Behaviours:** When satisficing, readers examined only the shelves geographically (and therefore topically) closest to the books they sought. In some cases the shelves were an active impediment to locating books; we observed two cases where the books sought were actually present on the shelves, but readers could not find them due to shelf layout or difficulties with the classification scheme.

**Implications for online browsing:** Satisficing only occurs when sought books are not present in the library. In this instance, nearest neighbours in search ranking may not support satisficing; in the example given above, the reader needed a textbook specifically; a workbook or popular science tome would not have met her needs. This implies that in addition to keyword similarity, these readers need to be able to examine semantic, non-bibliographic features of near-neighbour books—for example tone, length and presentation. Book surrogates presented by any system must include both semantic and bibliographic detail to support satisficing; this may be provided in the form of search refinement facets or tagging and folksonomies, for example.

Opportunism
Opportunism is browsing the shelves in an ad hoc manner when otherwise engaged in acquiring a pre-identified target book. Opportunism may or may not result in the borrowing of additional books.

**Motivations:** Opportunism occurs in a similar context to grab-and-go and satisficing: the initial intention is to borrow a single book or small number of books. Grab-and-go browsing leaves readers unmoved by the context in which the books they seek exist; satisficing leaves them negatively affected. In contrast, opportunism is a direct result of the shelves having a positive effect on readers; the shelves offer alternatives they had not intended to explore but that are welcomed. While it is possible for serendipity to be negatively experienced (Waugh, McKay, & Makri, 2017), in opportunistic browsing serendipity is both unexpected and positive. One reason given for such positive experiences was finding other books “for later boredom” (as one of the interview participants noted); another was the unplanned identification of further useful material. For readers who selected no further books, there seemed to be a satisfaction in knowing they had selected the right material for their needs; one commented that “the other books were too basic”. We observed 6 opportunists and interviewed a further two, meaning this group accounted for 17% of all participants.
**Behaviour:** Most readers who engaged in opportunism examined only a small area of shelving around their target material, however one systematically examined a whole call number subsection comprising 2-300 volumes. In at least one case a reader browsed two geographically distant areas of the library using this approach; she had identified two books on the history of Luxembourg using search, one in the history section and one in a geography section. The history section proved relatively fruitless for her needs, but she was surprised to find as much material as she did in the geography section. Of the readers who engaged in this type of activity, three left with only the books they had intended to take and the remaining six took additional books. This is slightly more than we would expect, given (for example) Beaulieu’s observations (Hancock-Beaulieu, 1993), suggesting that this approach to the shelves increases the likelihood that readers will borrow additional resources: these readers are ‘making their own luck’ (Makri et al., 2014).

**Implications for online browsing:** Facilitating opportunism in an online environment is challenging. It is possible that search results will offer enough options for readers that they explore and feel satisfied, however library log analysis suggests readers explore search results in a limited way (Cooper, 2001; Lau & Goh, 2006; Rowlands et al., 2008), and ebook borrowing records suggest the same (McKay, Smith, et al., 2015). While for some readers alternative options will be experienced as a distraction from their main task (Waugh et al., 2017), these readers welcome it while not anticipating it. These readers need more than an information ‘scent’ (Pirolli, 1997) to have the kind of rich information experience observed here; the shelves represent a full information experience. How best to provide an indication that such richness exists without limiting the experiences of other information seekers remains an open research question.

**Seeding by Search**

Seeding by search is the deliberate use of keyword or known-item search to generate a location index for an area of the shelves likely to contain relevant materials. This area is then browsed extensively.

**Motivation:** The intention of search seeding is to browse; search is only used as a means to an end. While this may look like traditional topical browsing, the means by which readers access a topic area on the shelves is relevant for how this behaviour might be supported online. Serendipity is an anticipated outcome of search seeding; the shelves are actively leveraged for what they have to offer. Search seeding is used to meet targeted but not highly specific information needs; one reader wanted to read a railway engineering manual for recreation on his lunchbreak (eschewing with disdain the biographies of railway engineers found nearby); another was looking for a modern book on android programming to upskill in his current employment role. We observed 6 of these readers and interviewed 4, 22% of all participants.

**Behaviour:** Searches may be for an individual known item, as with the railway enthusiast mentioned above who searched for a book he had read in the past; in contrast the android programmer searched merely for ‘android’. Search-seeded browsing may cover a smaller or larger area; while many readers engaged in search seeding examine only a small number of books, some cast their eyes over up to a thousand. Typical physical behaviours noted in observations include moving along the shelves, stepping away from and toward the shelves, removing and returning books and placemarking. All triage (determining whether to take a book) occurred at the shelves. These books were often compared on a range of axes, rather than simply on topic: for the programmer given above, the date the book was published was important. Another reader said during his process “I’ve got some generalist things, now I need something more specific”, meaning that his semantic interpretations of the books were important to him.
**Implications for online browsing:** Supporting search seeding online means ensuring readers can access and assess a large range of semantically related books quickly; search results presented ten by ten are simply not going to meet the needs of this group. They need to be able to see alternative books while they examine individual items to replicate the experience of triage at the shelves, and they need access to a range of semantic and descriptive information about the materials they are viewing. They need to be able to pan and zoom, but they also need to be able to placemark. Many of these features are not new—Marchionini identified some of them in the 90s (Marchionini, 1997) and Kwasnik identified others (Kwasnik, 1992). However, the requirement for visible representation of both candidate and potential candidate books simultaneously is new, as is the particular combination of these features to meet a specific browsing behavioural need.

**Seeded by location**

Location-seeded browsing is the return to a previously-known physical area of the library to browse the books again.

**Motivation:** Readers engaging in location seeding have similar information needs to those seeding by search: targeted but not highly specific. This approach is similar to the favourite locations mentioned in (Björneborn, 2008), but that work does not address the behaviour once readers reach these locations. Like search-seeded browsers, location-seeded browsers anticipate serendipity, however unlike search-seeded readers they have identified a fertile information environment for browsing without search. This browsing type accounted for 13% of all participants; 3 observations and 3 interviews

**Behaviour:** Most location seeding requires existing knowledge of the library to identify browsing locations, however we observed one reader who had “gotten a librarian to show them [where to browse]”. Location seeding is—except for conducting a search—broadly similar to search seeding.

**Implications for online browsing:** The main difference in requirements for online browsing systems between search seeding and location seeding is that in location seeding readers need to be able to return to a previously visited location. This might be in the form of persistent placemarking, or—in the case of those who accessed a librarian—overview.

**Wandering**

Wandering is using the whole library as a browsing ground in an attempt to find ‘a good book’.

**Motivation:** Wandering was rare; we witnessed only two readers, both at the State Library site (4% of all participants). It is the closest to the ‘General Purposive Browsing’ described by Apted (Apted & Choo, 1971), or Celoria’s recreational browsing (Celoria, 1969). These users had very loosely defined goals, expressed by one as “I have a few hours and I am looking for something interesting to read”. In contrast to Apted, they were not looking for information specifically, but rather meeting a need (“finding something to read”) that involved selecting information. For this type of reader, serendipity represents nearly the entire plan—they are looking to the library to inspire them.

**Behaviour:** Physical behaviour in wandering includes heavy use of subject endcaps to direct reader inspiration, considerable zooming, and scattershot browsing. They do not look at every book in a section or browse exhaustively—rather they hunt and peck. These readers both stood back from the shelves before grabbing books that seemed too far away to read. When asked why they chose a specific book, one of these readers commented that it was “a very good question”. Of all our types, this type of browsing looks the most like browsing described by Bates (Bates, 2007).
Implications for online browsing: Supporting online wandering requires an overview of the library, excellent support for zooming, and an ability to see a wide range of books and select those that look interesting from a distance; book covers are a known element of book decision making, for example (Ooi, 2008; Saarinen & Vakkari, 2013). These readers are not supported at all by traditional keyword search; even where they can generate a keyword, the limited presentation of results does not meet their information seeking needs at all.

Summary
We have presented six behavioural strategies observed in the library shelves. Of these, one is known-item searching, and the others include at least some measure of browsing. The types are differentiated by the motivations, intentions and behaviours of the readers engaging in them, and each type has different implications for online browsing. While the types have been presented here in relation to search, there are a number of other features that describe each type; these are presented below.

Additional Features of the Six Types
In addition to distance from search, the types can be classified in three ways: The influence of serendipity on the reader’s experience (and vice versa); how precise an information need the reader had when they approached the shelves; and how broad (or narrow) their starting point for browsing was.

Serendipity is a feature frequently described as valuable by the users of library shelves (Makri et al., 2007; Stelmaszewska & Blandford, 2004), and indeed previous research suggested that library shelves are almost perfectly designed as serendipity engines (Kleiner et al., 2013). In our browsing types serendipity can be experienced as planned (much like the creatives making their own luck in (Makri et al., 2014)), grudging (where it is utilised in contrast to the original plan of the user) or welcome but unanticipated.

The browsing we observed was often predicated on how specific the original information need was. In some cases, our users had ‘ill-defined’ needs that could not be readily described as queries (Belkin et al., 1982), or needs that were better met by recognition than specification of objects (Marchionini, 1997). While our observations are a snapshot of what might be part of a broader information seeking process, we did not witness information seekers using browsing to refine their needs, as described in (Kuhlthau, 1991; Kwasnik, 1992; Marchionini, 1997; Rice et al., 2001); each encounter focused on a relatively fixed information gap.

A further classification of these types, and one key to the construction of supportive information systems, is the presence or absence of search. While search and browse can be interleaved as described in the major information seeking models (Kuhlthau, 1991; Marchionini, 1997), they are not always, and online browsing needs to support all five approaches presented here to be successful.

The final classifier on which we can describe our browsing types is whether an original sought item was present, absent or unspecified. Marchionini would describe this as a feature of the information environment (Marchionini, 1997), and for our users it was a key defining feature: the absence of a sought item could lead to two further behavioural types.

The types are described according to these classifications below:

<table>
<thead>
<tr>
<th>Type</th>
<th>Serendipity</th>
<th>Information need</th>
<th>Search</th>
<th>Item Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grab-and-go</td>
<td>Unwanted</td>
<td>Specific</td>
<td>Present</td>
<td>Present</td>
</tr>
</tbody>
</table>
Table 1: Browsing types in terms of serendipity, information need, search and book availability

<table>
<thead>
<tr>
<th>Activity Type</th>
<th>Serendipity</th>
<th>Information Need</th>
<th>Search</th>
<th>Book Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisficing</td>
<td>Unwanted</td>
<td>Specific</td>
<td>Present</td>
<td>Absent</td>
</tr>
<tr>
<td>Opportunistic browsing</td>
<td>Unexpected</td>
<td>Specific</td>
<td>Present</td>
<td>Present</td>
</tr>
<tr>
<td>Search seeding</td>
<td>Anticipated</td>
<td>Loosely defined</td>
<td>Present</td>
<td>N/A</td>
</tr>
<tr>
<td>Location seeding</td>
<td>Anticipated</td>
<td>Loosely defined</td>
<td>Absent</td>
<td>N/A</td>
</tr>
<tr>
<td>Wandering</td>
<td>Anticipated</td>
<td>Very loosely defined</td>
<td>Absent</td>
<td>N/A</td>
</tr>
</tbody>
</table>

As we can see, many of the elements predicted by previous research are present—some browsing episodes contain unanticipated serendipity, and some are about readers actively leveraging serendipity (Makri et al., 2014). Some episodes are about dealing with a specific unmet need; others are about meeting needs that are so amorphous as to defy description (Kuhlthau, 1991; Marchionini, 1997). Some approaches use search, some don’t. The information environment is important, both in terms of whether an item is available, but also how serendipity and discovery are afforded (Marchionini, 1997). While each of these features individually could have been predicted from the literature, the way in which they combine to describe this browsing typology is entirely novel, and each feature has predictive value for adequate support in online book browsing. To support ebook browsing, we must support each element of the typology presented here.

Discussion

In this section we first describe our learnings about the nature of browsing in relation to the existing literature, then examine how these learnings might be applied to improve the experience of online browsing.

On the Nature of Browsing

While one contribution of this paper is to divide browsing up into six activity types, this work also reflects some truths about library shelf browsing that apply to all the types we describe except grab-and-go. Browsing at the library shelves isn’t just about the sequential examination of interesting things as (Bates, 2007) suggests. Nor can we rely on Ellis’ circular definition of good browsing systems, that they ‘allow the searcher to browse, while the system provides the searcher with the information he [sic] requires to browse effectively’—where the information Ellis describes is primarily bibliographic information (Ellis, 1989).

Shelf browsing is an exploration not just of individual objects, as Bates suggests, nor of individual facets of objects as Ellis’ model implies. Instead, we posit that shelf browsing is an exploration of information objects in the context of their relationship to each other. On subject-classified shelves, the predominant relationship is topic, author-sorted shelves focus on authors. These relationships do not operate, however, in a single dimension: even when subject browsing, readers are using multiple facets of any given object to make decisions about its value: these facets may include (without ever touching the shelf) cover style, size (and implicitly, length), age, and use by previous readers (Blandford et al., 2006; Hinze et al., 2012; Makri et al., 2007; Mikkonen & Vakkari, 2012; Stelmaszewska & Blandford, 2004). In our observations we noted readers exploiting date relationships; topical relationships; book type relationships (e.g. textbooks versus workbooks); relationships of title and keyword. Indeed, the relationships between items drive every type of browsing in our typology: satisficing is driven by a need for the topically and semantically similar; opportunism is driven by the serendipitous presentation of the topically similar; search- and location-seeded browsing are driven by an understanding that shared physical location also implies shared topic; and wandering is characterised by a process of narrowing down objects based on their relationships. If we fail to recognise the need to examine these relationships and exploit them when
we design information systems, we will surely fail to offer readers the browsing experience they need.

When we compare our typology to the other frameworks described at the beginning of this paper (Apted & Choo, 1971; Celoria, 1969; Herner, 1970; Rice et al., 2001), we see a number of similarities: Our grab-and-go browsers are those looking for a specific item in Rice, and Herner’s directed browsers. Our wanderers are represented in every other taxonomy (perhaps surprisingly, since they appear least frequently). Rice refers to this behaviour as incidental, Herner as undirected. This behaviour is similar to Apted’s general purposive browsing, and an exact match for Celoria’s recreational browsing: it seems this type of browsing activity is the easiest to recognise. In between these types, though, each model represents only a few of our types. Rice’s description of looking for items that share some characteristics could apply to our location and search-seeded browsing, where readers look for topically related materials; these types also match Celoria’s ‘professional browsing’, and Herner’s semidirected browsing. None of the previous models has accounted for satisficing or opportunism nor have they expressed the underpinning behaviours involved in seeded browsing of either type. While these behaviours are well known to practicing librarians, their absence from browsing typologies means that they have not been considered in relation to other behaviours, nor have they been considered in designing online systems. This paper aims to integrate these behaviours into a general understanding of browsing, and to make them visible to designers of online systems.

How to support online browsing

Previous work on online information seeking generally, and browsing specifically, has consistently noted that browsing is poorly supported online (Ellis, 1989; Kleiner et al., 2013; Marchionini, 1997; Thudt et al., 2012). How best to address this has been considered by a number of researchers: Ellis (Ellis, 1989) demanded simply that any metadata held be made browsable (though he was generally referring to providing clickable scrolling lists); even this low bar is not met by many information systems. In contrast Wilson suggests that actually, human computer interaction cannot generate requirements for information systems, just recommendations (Wilson, 1999), and Martin et al. (Martin, Greenspan, & Quan-Haase, 2017) suggest that attempting to move away from the physical library entirely is a mistake in terms of serendipity. To abandon any attempt to specify online browsing in the face of an ebook revolution, though, is irresponsible to future generations of readers. When over half of a library collection is online and much library use occurs in physical locations away from the library, supporting digital browsing has become a necessity.

Marchionini suggested a number of required features of online browsing systems, including highly detailed and interactive representations of information and the means by which to manipulate them, support for zooming and panning, support for more and less detailed user probes to bring back useful information hunting grounds, and ways of limiting the information space such that the information is both manageable and relevant (Marchionini, 1997). Kwasnik also noted a requirement for zooming, and further noted that placemarking must be supported (Kwasnik, 1992). More recently Whitelaw noted the need for overview in complex information tasks, and for that overview to be carefully considered and as neutral as possible (Whitelaw, 2015).

Overall, many of the needs we have identified in our typology—zooming, overview, placemarking, bringing geographically distant but semantically related items together and accessing semantic information alongside bibliographic information—are not new. We have added a further two needs—being able to examine one item while retaining a sense of one’s other choices, and supporting (particularly opportunistic) readers in understanding that there is more available if they
want it. What is new is the identification of all of these components as key elements of library shelf browsing, and as having relevance to specific types of browsing. We have further extended the literature on what browsing entails, using empirical evidence to suggest that what has previously been described as ‘browsing’, or at most grouped into four tasks, is actually at least six information seeking trajectories, each with specific associated needs. Our options for meeting these needs are increasing with, for example, touch technology, but how best to leverage these technologies to support each browsing activity identified here remains future work.

Conclusions
The revolutionary shift from print books to ebooks in libraries has made the need for better online browsing pressing. Some readers will not use ebooks due to the lack of a browsing facility; those who will still use ebooks lose a fundamental part of the human information seeking process, along with a major facilitator of serendipity.

It is widely acknowledged that there is need for significant research into how to provide better online browsing systems; for such an important component of information seeking, however, browsing has consistently been treated as a ‘black box’ activity. While a number of papers have examined library shelf use, none has been designed to understand browsing, or how to facilitate it online. Equally, some authors have proposed typologies of browsing behaviour, but none since 1999, and none with a clear grounding in empirical evidence. It is this gap we address in this paper: we present an empirically grounded typology of browsing behaviour, along with recommendations for how these behaviours might be supported online. The behaviours in our typology are influenced by information seeker needs, the information environment, and information seekers readiness or otherwise to experience serendipity. By understanding these browsing types, we have identified a number of key requirements for online book browsing, including zooming, placemarking, and the ability to examine one item while retaining a sense of the other options available.

We have further refined our understanding of what browsing means, at least where library shelves are concerned: book browsing specifically relies on a combination of the individual items Bates speaks of (Bates, 2007), and the relationships implied by Ellis’ discussions of browsing support (Ellis, 1989).

There are many questions raised by this work, including the extent to which personal preference influences selection of browsing type, how often information seekers switch from one browsing type to another mid task, and what a browsing interface to support all these approaches might look like. These questions remain future work.

Acknowledgements
Our thanks to the State Library of Victoria and Swinburne University of Technology Library for their support in data collection. We further thank Rebecca Parker for her invaluable comments on a draft.

References


Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
McKay, D; Chang, S; Smith, W; Buchanan, G

Title:
The Things We Talk About When We Talk About Browsing: An Empirical Typology of Library Browsing Behavior

Date:
2019-12-01

Citation:

Persistent Link:
http://hdl.handle.net/11343/241941

File Description:
Accepted version