ANNA MIKKONEN
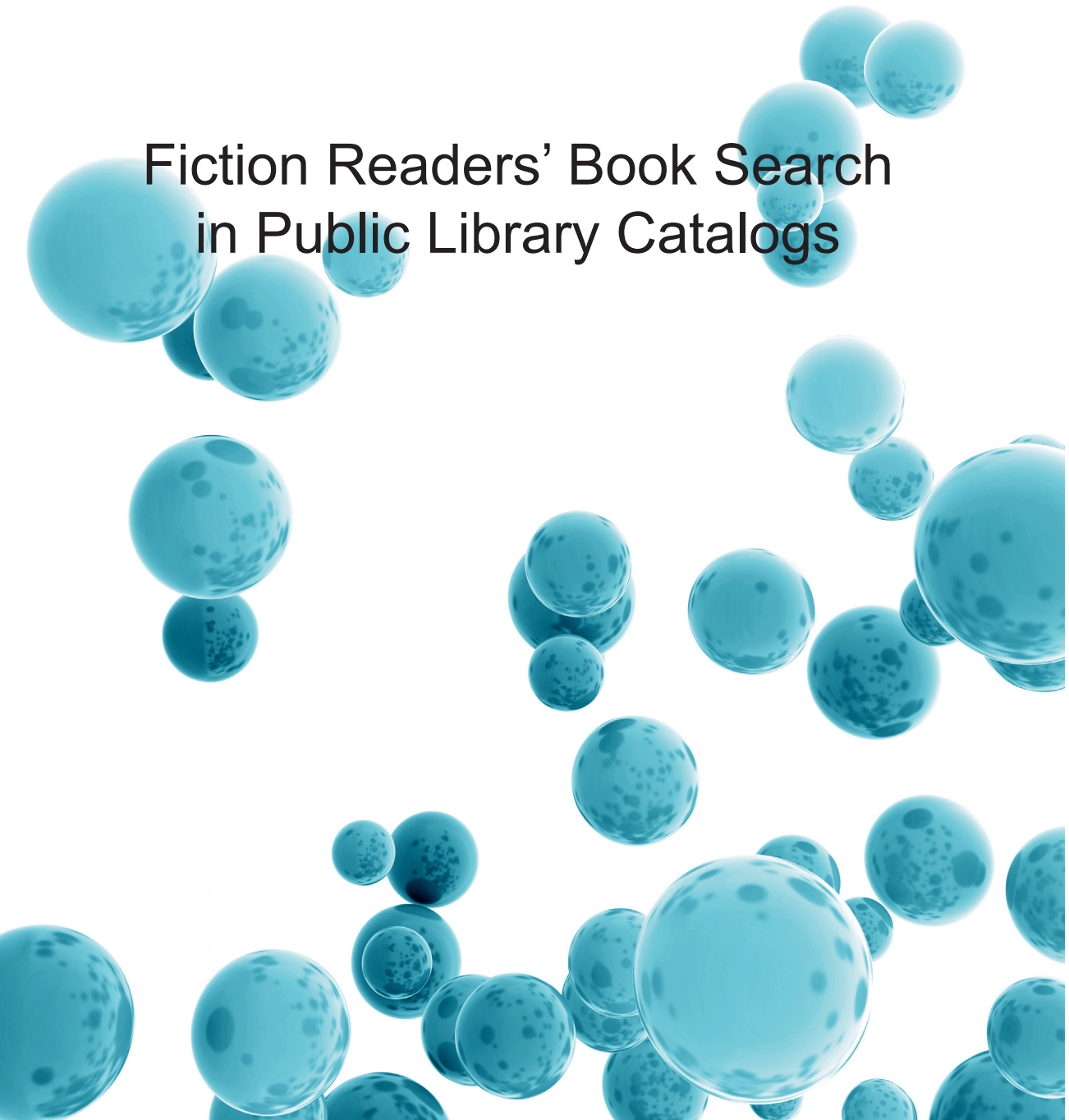
Fiction Readers' Book Search in Public Library Catalogs

ANNA MIKKONEN

Fiction Readers' Book Search
in Public Library Catalogs

■

UNIVERSITY OF TAMPERE

ANNA MIKKONEN

Fiction Readers' Book Search
in Public Library Catalogs

Cover design by
Mikko Reinikka

ACKNOWLEDGEMENTS

encouraging presence in my life as my godmothers. My friends Hannele, Karoliina, Jenni, Pia, Miia, Marianne, Tuuli, Anna-Marika, Anna R. and Anna N., thank you for enlightening this journey. It would have not been the same without you.

My brother Mikko, thank you for your support and for being a great example of an open-minded adventurer. My nephew Oiva, thank you for our happy moments together.

My parents Teija and Heikki, you have taught me the enthusiasm for literature and knowledge. You have always let me find my own path, supporting me in my decisions. I am grateful for your love and support.

My husband Sami, it has been such a journey from the corridors of nursing school to the work life, academia, marriage and parenthood. I have not been the easiest and smoothest companion with my ambitions. Thank you for unquestionably believing in me and us.

My children Vilja and Väinö, I feel privileged to be a mother for two little people with such a unique spirit and an independent mind. I wish you never lost them. And always remember: *"I do believe something very magical can happen when you read a good book".* -J.K. Rowling-

Hämeenlinna, June 17th 2017

Anna Mikkonen

ABSTRACT

The study investigates fiction readers' book searching in public library catalogs. The goals are to describe the fiction book search process in an interaction with information systems, to reveal typical search approaches for fiction books in digital environments, to specify the understanding about exploratory browsing by eliciting the navigational patterns of fiction readers, to evaluate fiction readers' interest criteria for novels, and to describe the association of various preference patterns for fiction with the book selection process. The study makes use of two separate data sets: a survey data collected in the year 2010 for studying the outcomes of public libraries, and a data from a user-study conducted in the year 2012 for studying the search behaviors of fiction readers in online environments. In order to determine whether various background variables and demographic factors are associated with the book search behavior in public libraries, the survey data is analyzed. For understanding fiction readers' search behaviors in digital environments, a data from a controlled user-study with 80 genuine fiction readers working on various search scenarios for fiction is analyzed. The search scenarios are designed to reflect the following search tactics: known-author search, topical search, open-ended browsing, search by analogy, and searching without a query. As a baseline system, a traditional online library catalog Satakirjastot is used. As an enriched test system, the web service for fiction BookSampo is used.

The study enlightens readers' typical search approaches for books in various search scenarios and describes the differences in user behavior between different library catalogs. The study reveals that the fiction book search shares characteristics both with casual-leisure and work-based search. The most often used search tactics for accessing books in public libraries are the known item search together with the browsing search. In public library catalogs, the most used search moves are querying, search results inspection, and book page examination. The results show that readers' search behaviors differ between the two library catalogs, but result in an equal search success. The results demonstrate that browsing task type is associated to readers' document viewing behavior in terms of observed search result pages, opened book pages and dwell time on book pages. When browsing for topical novels, most effort is required to select somewhat interesting novels. When browsing for good novels,

most effort is required to select very interesting ones. The findings show a context-related pattern in readers' fiction book selections. A combination of readers' search capacities, "behind the eyes" knowledge, affective factors and a well-functioning interaction with a system used results in a successful book selection. The results demonstrate that various preference patterns for fiction are associated to readers' search behavior, i.e. the number of viewed search result pages, opened book pages, dwell time on book pages and the type of search queries. Based on the findings, three search tactics for fiction books in library catalogs are presented: 1. Focused querying, 2. Topical browsing, and 3. Similarity based tactic. The most popular search tactic in each search scenario is "focused querying" with known author.

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

The dissertation consists of a summary and the following original research publications:

I. Mikkonen, A. & Vakkari, P. (2012). Readers' Search Strategies for Accessing Books in Public Libraries. Proceedings of Information Interaction in Context 2012, ACM, New York, NY, pp. 214-223. doi: 10.1145/2362724.2362760

II. Mikkonen, A. & Vakkari, P. (2016). Finding Fiction: Search Moves and Success in Two Online Catalogs. Library and Information Science Research 38(1), 60-68. doi: 10.1016/j.lisr.2016.01.006

III. Mikkonen, A. & Vakkari, P. (2015). Books' Interest Grading and Fiction Readers' Search Actions During Query Reformulation Intervals. Proceedings of Joint Conference on Digital Libraries 2015 in Knoxville, Tennessee, ACM, New York, NY, pp. 27-36. doi: 10.1145/2756406.2756922

IV.   Mikkonen, A. & Vakkari, P. (2017). Reader Characteristics, Behavior and Success in Fiction Book Search. [Accepted for publication in Journal of the Association for Information Science and Technology]. doi: 10.1002/jasist.23843

V.   Mikkonen, A. & Vakkari, P. (2016). Readers' Interest Criteria in Fiction Book Search in Library Catalogs. Journal of Documentation 72 (4), 696-715. doi: 10.1108/JDOC-11-2015-0142

These publications are referred to as studies I–V in the remainder of the dissertation summary.

The first author collected the data except the first study, performed the data analysis, and was in charge of writing the articles. Professor Pertti Vakkari supervised and commented on the work throughout the research process.

1   INTRODUCTION

While the Internet and modern technologies are substantially changing people's reading behavior (Bergstrom & Hoglund 2014; Shimray et al. 2015; Zhang & Kudva 2014), the book still remains the medium of choice for millions of people for leisure reading (Zhang & Kudva 2014). The benefit of pleasure reading is associated to *"coping with the world through simulated situations"* (Gold 1990). According to Wittgenstein, *"the limits of my language means the limits of my world."* Thus, we read to expand our language; in imaginary literature we seek clues to recognize ourselves, to tell us what we think and how we perceive the world. The present study focuses on investigating fiction readers' book searching in an experimental setting, proceeding from an idea that different literary preferences have an association with the readers' book selection process in an information system. Previous studies (Buchanan 2011; Goodall 1989; Saarinen & Vakkari 2013; Spiller 1980) on fiction book searching have focused on examining the book selection as it occurs at the shelves of public libraries or in the bookshops. However, little is known about how readers currently search for recreational reading materials in digital environments (Adkins & Bossaller 2007; Elsweiler et al. 2011). The study contributes to the comprehensive understanding about search behaviors for casual-leisure reading material. It reveals factors that are associated with search success when searching for novels in digital environments, which is essential as digital fiction collections continue to grow. It enlightens the differences and similarities between a search for recreational reading material and search for non-fiction information resources, and contributes to the research methods applied for examining casual leisure information behavior.

Fiction readers make up a large portion of the public library clientele as a) leisure reading focuses on fiction and b) fiction is often searched for and found in public libraries (Maker 2008; Moyer & Weech 2005; National Endowment for the Arts 2012; The Reading Agency 2013; Van Riel, Fowler, & Downes 2008; Yu & O'Brien 1996). According to Mehtonen (2016), recent developments in information technologies have profoundly affected our attitudes towards the function of public libraries as their operational environment has been digitalized to the great extent. At present, public library online catalogs are actively developed to meet the needs of diverse user groups. The concept of *a metadata-enriched library catalog* (Eden 2002) has

arisen to refer to the various metadata elements providing diverse access points to library collections. Instead of asking a librarian, hints of interesting books are searched from the user generated content and enriched metadata in online library catalogs. In spite of the recent developments in online library catalogs, little is still known about fiction readers' search behavior in library catalogs while discovering books to read. Here, we take an evaluative approach to the fiction book search process with an attempt to assess the functionality of the metadata-enriched library catalog from a user-perspective.

Research on book search behavior in digital systems has mainly focused on retrieving non-fiction (e.g. Hessel & Fransen 2012; Lau & Goh 2006; Moulaison 2008; Waller 2010). However, searching for fiction differs from the retrieving of non-fiction as the search process is not always triggered by an explicit information need to be solved. Instead, it begins often with a serendipitous need to find *"just good books to read"*, which engages fiction readers in exploratory search behavior (Elsweiler et al. 2011; Pejtersen 1989; Wilson & Elsweiler 2010). The scope of the present study includes examination of fiction readers' search behaviors in two online library catalogs in various search scenarios. The aims are, firstly, to describe the fiction book search process in information systems, and secondly, to elaborate the differences between a traditional and an enriched public library catalog. Moreover, the aim is to illuminate the differences and similarities between a fiction and non-fiction search, and to contribute to the research methods applied for examining casual leisure information behavior. The study is a compound dissertation consisting of a summary and five papers. Together the five papers contribute to the understanding of fiction readers' search behavior by answering the following research questions:

1. What are the typical search approaches for fiction books in public libraries? (Study I)

2. What are the typical search moves for fiction books in online library catalogs and how are they associated with search success in different search scenarios? (Study II)

3. How are readers' document viewing behavior and relevance assessments associated with fiction search success? (Study III)

4. How are various preference patterns for fiction books associated with fiction book search and search success in library catalogs? (Study IV)

5. What are the navigational patterns of fiction readers while selecting novels in library catalogs? (Study IV)

6. What are the interest criteria for fiction books applied by fiction readers in library catalogs,? (Study V)

These questions present the general focus of the research presented in this thesis. In each study (I–V), several more precise research questions are found.

The structure of the thesis is as follows: Study I surveys the major search tactics applied by readers in public libraries and investigates the association of demographic factors and reading activity to readers' search tactic selection. Study II examines the association of individual search moves and search success in two different online catalogs, and study III focuses on investigating fiction readers' document viewing behavior and novels' interest criteria during query reformulation intervals. Study IV is about various preference patterns for fiction reading and their association to readers' search behavior and search success. Study V presents a categorization of interest criteria for novels based on fiction readers' interpretations of the influential factors in fiction book selection in different search scenarios.

The thesis has a few essential concepts at its core. The thesis focuses on fiction readers' *search behavior* in online library catalogs. *Search behavior* refers to how users search information on the Internet in general. Search behavior constitutes of *search actions* (*search moves*). *A search action* is an identifiable basic action (i.e. conducting a keyword search, a click to view a search result page or a book page) during the search process (Bates 1990). In fiction book search, the major search actions are issuing *a search query*, inspecting *a search result page* and examining *a book page*. *A search query* is a formulation of a precise request for information in the form required by the information retrieval technique (Ingwersen & Järvelin 2005). *A search result page* (SERP) is the page displayed by a search engine in response to a query by a searcher. *A book page* (a title page) is the page that displays the publication information and the metadata for a particular item on an online library catalog. While selecting books in online library catalogs, readers alter the use of search actions flexibly according to a search scenario. Sequences of search actions that are temporally and semantically related are named as *search tactics* (Wildemuth 2004). What is characteristic of fiction book search is *browsing*, which refers to intentional searching without a specific goal that depends on recognizing interesting information. In browsing for novels in a library catalog, a reader is navigating through an online environment indirectly with an aim to encounter interesting novels. In the present study, the fiction book search is examined in an experimental setting during various simulated search tasks. *A simulated search task* is a specific scenario, a cover story designed for a research setting that aims to trigger individual information problems in order to test users in a controlled manner (Ingwersen & Järvelin 2005).

After the introductory chapter, the second chapter of the summary presents different approaches to fiction book search and describes the relevant previous

research. In the third chapter, the data, the participants and the experimental procedures are described. Chapter 4 presents the findings, and Chapter 5 discusses them. Finally, conclusions are presented in Chapter 6.

2   APPROACHES TO FICTION BOOK SEARCH

This chapter outlines the framework for the present study (Figure 1) and describes previous empirical studies regarding fiction reading and book search behavior in public libraries. The chapter presents studies of various motives for fiction reading and reviews fiction book search as it occurs at the shelves of public libraries. It describes fiction book search in digital environments and as a casual leisure activity.
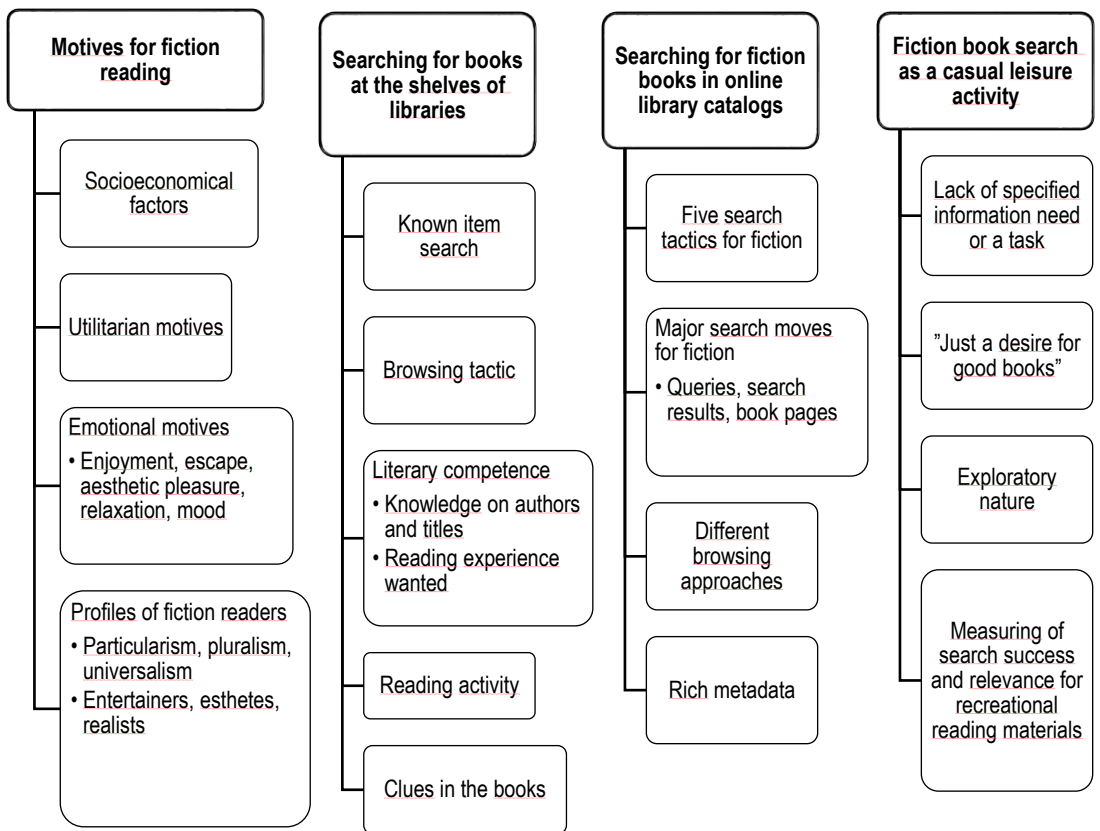
Figure 1.   The theoretical background of the study.

2.1   Motives for fiction reading

According to Guthrie & Wigfield (2000), a reading motivation can be defined as *"the individual's personal goals, values, and beliefs with regard to the topics, processes and outcomes of reading".* The aim of this section is to describe the motivation for fiction reading and to distinguish various fiction reading preferences that can be considered having an association with fiction readers' search behavior and the book selection process. Studies describing socioeconomic characteristics affecting the tendency to read fiction and studies on various literary preferences are reviewed.

Characteristics affecting the tendency to read fiction have been identified in several studies. In studies, the fiction reader has been approached mostly through demographical features building the classifications of readers on the basis of their age, gender, occupation, etc. (Yu & O'Brien 1999). It has been shown in several studies (i.e. Griswold et al. 2005; Spiller 1980; Jennings & Sear 1986; Moyer 2008; Ross 2001) that women are more likely to read fiction and more likely to pursue leisure reading as a hobby than men. Tepper (2000) found that after controlling education level, women's likelihood to have read a fiction book in the prior 12 months was over 2 times greater in comparison with men. Women are also more likely to be heavy readers than males and they tend to read a greater variety of books than men (Ross 2001; Tepper 2000). According to Tepper (1998), it appears that differences in gender role stereotypes and childhood socialization offer the strongest explanation for the gender gap in fiction reading, as in reflections for appropriate leisure activities for boys and girls. It has also been shown (i.e. Griswold et al.; 2005; Ross 2001) that recreational reading declines with age so that heavy readers are more likely to be younger rather than older.

In addition to gender and age, education has a major impact on leisure reading as highly educated people read more than people with lower education (Scales & Rhee 2001; Statistics Finland 2009). Education is also positively associated with the frequency of reading fiction and the type of fiction read (Ross et al. 2006; Kraaykamp & Dijkstra 1999). Eskola (1979) found out that the uses of literature were more varied among people with higher education than in people with low education. Eskola (1979) suggested that the ability of different types of literature to offer inspiration to read varies between different socioeconomic groups.

Few psychological and sociological studies focusing on fiction readers' different reading preferences have been carried out. Kraaykamp & Dijkstra (1999) examined

social differentiation in book reading preferences using a Dutch nationwide survey. Findings indicate that educational resources were important in explaining reading preferences as a higher educational level and the stimulation of literary reading in secondary school determined reading preferences for complex books. Also, social motives were meaningful to book reading preferences as an attitude that acknowledged the universal value of literature led to a preference for complex and prestigious books (Kraaykamp & Dijkstra 1999.) Stockmans (2003) conducted a survey which sought to understand whether psychological factors, such as attitudes, motivations and ability, could explain preferences for a reading material of varying complexity and negatively affect the explanatory power of social class. The results indicated that reading attitude as well as motivation and ability to read complex fiction varied within social classes and had an effect on the tendency to read complex fiction independently of the effect of social class (Stockmans 2003.)

In various studies (i.e. Gerrig 1993; Krashen 1993; O'Rourke 1993; Horsman 1990; McClellan 1981; Usherwood & Toyne 2002) it has been confirmed that reading imaginary literature serves to satisfy a wide variety of needs that can roughly be categorized as utilitarian (Maryl 2008; Miesen 2003; Moyer 2007; Usherwood & Toyne 2002; Scherf 1990) and affective (Maryl 2008; Ooi & Liew 2011; Radway 1984; Ross 2001; Usherwood & Toyne 2002). Reading to fulfill utilitarian needs functions as a practical motive towards pleasure reading and reflects the development of practical knowledge through reading imaginative literature (Maryl 2008; Miesen 2003) whereas the affective dimension of reading is related to aspects such as enjoyment and entertainment (Miesen 2003; Ooi & Liew 2011).

The outcomes of pleasure reading were studied by Moyer (2007) who used a combination of surveys and interviews to explore the relationship between the educational and recreational outcomes of leisure reading in a context of incidental information acquisition. The results show that educational outcomes gained through fiction reading concerned learning about other times, places, cultures, people and relationships and offered enrichment and different perspectives (Moyer 2007). Similarly, Maryl (2008) found that pleasure reading offered guidance in unfamiliar situations and its role in acquiring new, practical knowledge was important for pleasure readers. The utilitarian aspect of pleasure reading was also highlighted by Usherwood & Toyne (2002), who noticed that readers often reflected how imaginative literature contributed to their practical knowledge and learning of the world, especially of other cultures and customs. Pleasure reading instructed readers also in relation to their own personal development, as reading about other people's experiences was seen to ratify their own. Reading improved readers' ability to relate

to other people too (Usherwood & Toyne 2002.) Thus, engaging in the simulative experiences of fiction literature may increase the understanding about others and enhance the capacity for empathy (Mar & Oatley 2008).

Ooi & Liew (2011) asked adult fiction readers about their motives and attitudes towards fiction reading. The results show that fiction readers enjoyed the *entertainment, escape* and *comfort* that reading as a pleasurable activity offered. Similarly Maryl (2008) found that processing of a literary text relies heavily on the *reader's mood* and their need to alter or deepen it. The emotional motives that guided the fiction reading process found by Maryl (2008) were *seeking specific emotions, emotional stimulation, relaxation and escape*. Also Radway (1984) distinguished the aspects of the emotional impact of pleasure reading in everyday life: escape from daily problems and relieving tensions. Usherwood & Toyne (2002) showed that *escapism* was the most conscious perception that readers derived from the act of pleasure reading. For readers, escapism meant actively being involved in other worlds, but also escape through *the aesthetic pleasure* that the works of literature offered. Together with escapism the most frequently cited emotional motivations for pleasure reading were *relaxation* and *release from daily pressures* (Usherwood & Toyne 2002.)

The attitude towards adults' leisure reading was studied by Miesen (2003) who measured adults' reading motives with a questionnaire containing behavioral beliefs towards for literary reading. The study revealed five major categories of motives for fiction reading:

Affect-Enjoyment, where the affective and experiential aspects of literature were emphasized

Utility-Intellectual development, which reflected the idea that reading literary novels is instrumental in obtaining intellectual enrichment

Prestige-Self-cohesion, where the items were about gaining prestige, reading to structure one's life and providing oneself with an identity

Utility-Broadening one's horizon, which reflects the idea that reading is instrumental in obtaining a broader view on life

Relief from boredom, where the items reflected the idea that reading literature is instrumental in chasing away boredom and killing time.

The study revealed that reading literary novels had a positive image when it came to fulfilling affective needs, such as enjoyment and entertainment. Utilitarian needs were also addressed. Utilitarian needs were explained reflecting the idea that reading literary novels was instrumental in obtaining intellectual enrichment and development. Thus, the study suggested that alongside the enjoyment that appears in the reading process, the utilitarian outcomes appear to be equally important

(Miesen 2003.) Miesen's findings appear similar to Ross's (2001). Ross interviewed 194 avid readers to find out how they chose books for pleasure reading. The findings suggested that the affective dimension was involved in a book selection from choosing a book according to mood to valuing a book for its emotional support in providing confirmation, courage or self-acceptance (Ross 2001). Furthermore, readers reported some educational outcomes of fiction reading, the most common being awakenings, new perspectives and the expansion of possibilities.

In addition to various motives of pleasure reading, previous empirical research has revealed various profiles of fiction readers. Yu & O'Brien (1999) studied 221 readers' reading habits. They distinguished six types of fiction readers, which differ from each other in the following dimensions: the frequencies of borrowing fiction, the number of authors currently read, the approaches applied to searching for fiction, literary/recreational orientation (i.e. preference for books on the serious fiction/light fiction continuum), varieties of books the reader reads and confidence in selecting books. By combining and reducing these dimensions, Yu & O'Brien (1999) created a typology of fiction borrowers. Depending on the variation of the authors and types of literature read, the readers' reading orientation was called particularism, pluralism or universalism. Particularists' reading scope was narrow as they read a very small number of particular authors, pluralists' had a well-defined reading interest where they read several authors and universalists' reading interest was diverse as they chose books whatever looked appealing (Yu & O'Brien 1999.)

Lukin (1994) identified three types of fiction readers when studying readers' ways of searching for interesting novels in a public library. The central distinguishing factor between different reader types was readers' expectations of a good novel. Lukin distinguished between entertainers, esthetes and realists. Entertainers were seeking relaxation and distraction from daily routines via recreational reading. They read mostly entertainment literature, and their orientation towards reading was mostly recreational. For entertainers, the subject of the novel and easy readability were important factors. Entertainers expected that the plot of the novel would carry along, and the characters and actions were fascinating (Lukin 1994.) Esthetes were looking for the aesthetic pleasure that the works of literature offered. The artistic and esthetic experience produced by a novel was the major criterion of a good novel for the esthetes. It consisted of an artistic composition of the text, of language use and of profound worldview (Lukin 1994). Realists read for deepening their knowledge about certain subjects. Realists preferred especially realistic texts and historical novels (Lukin 1994).

Lukin's (1994) classification was used also by Saarinen & Vakkari (2013) who studied fiction readers' tactics for book selection and examined the attributes of books that readers perceived as indicators of good novels. The findings confirmed Lukin's (1994) distinction: entertainers read mostly fiction belonging to the genres of entertainment novels; esthetes read high standard novels and realists preferred novels which taught them about the world and the society. Entertainers were seeking relaxation and distraction from daily routines from pleasure reading. Entertainers expected that the plot of the novel would carry along, and the characters and actions would be fascinating. Esthetes were looking for the aesthetic pleasure that the works of literature offered. The artistic and esthetic experience produced by a novel was the major criterion of a good novel for the esthetes. Realists were reading to deepen their knowledge about certain subjects. Realists preferred especially realistic texts and historical novels.

Summary of motives for fiction reading. To sum up, previous studies have revealed a wide variety of motives for reading imaginary literature. An affective dimension is involved in the book selection from choosing a book according to mood to valuing a book for its emotional support (Ross 2001). Alongside the enjoyment of the reading process, utilitarian outcomes such as reading fiction for learning and practical knowledge appear to be equally important (Miesen 2003; Usherwood & Toyne 2002). In addition, escapism has been named as the most conscious perception that readers derive from the act of pleasure reading. For readers, escapism means actively being involved in other worlds, but also escape through the aesthetic pleasure that the works of literature offer (Usherwood & Toyne 2002.) The attractiveness of a particular book is related to the motivation of reading and to the goals a reader wishes to gain through reading a particular book (Usherwood & Toyne 2002). The motivation and frequency of pleasure reading, the most preferred authors and types of literature, together with the attributes that readers perceive interesting in fiction books, distinguish readers into various reader types.

2.2   Searching for books at the shelves of libraries

The aim of this section is to describe empirical studies on search behavior and search tactics for fiction books inside public libraries. *A search tactic* is here defined as a practical book selection method that the reader uses to encounter desired books (Slone 2000; Yu & O'Brien 1996). Previous empirical studies (Spiller 1980; Yu & O'Brien 1996) on readers' fiction book selection in public libraries have revealed that the methods readers apply in their fiction searching and selection can roughly be broken down into two types, which are searching for a specific author or title and browsing. Searching for a specific item requires a reader to know in advance the author's name or the title of the book. According to Yu & O'Brien (1996), searching for specific books relies as a major part on authors' names. Known item search is a search strategy which is most supported by current library catalogs in public libraries. Readers perceive it easy to check the availability and location of a certain novel in the library catalog (Saarinen & Vakkari 2013; Slone 2000).

Compared to a known item search, browsing, or an unknown item search, is a more complex search strategy where previous reading experiences, elements from the books themselves, affective and personal factors and the informational needs of a person are present at the same time (Ross 2001). Moreover, browsing is characterized by serendipity which refers to an unexpected discovery of useful or interesting information. Serendipitous browsing as a method of book selection involves intentional but undirected "looking around" in public libraries and skimming the shelves, book displays, returned loans and book exhibitions with the reader hoping to encounter interesting books (Goodall 1989). A combination of browsing and known item search has been found to be a very popular approach to fiction book search in public libraries (Spiller 1980). A combination of a known item search and browsing most likely occurs when a) a reader begins with an author name and when failing to get the desired books, wounds up with other methods, or b) a reader begins to browse but ends up selecting books based on the recognition of author names (Goodall 1989; Jennings & Sear 1986). In addition to a known item search and browsing, Yu & O'Brien (1999) present *the category search*, where readers are looking for a particular category of books like genre and try to find books within that category.

In addition to revealing typical search approaches for fiction books in public libraries, previous research (i.e. Ross 2001; Ooi & Liew 2011; Saarinen & Vakkari

2013) has identified various factors associated with the use of these search tactics. These factors include reading activity and the type of books most read, which have been shown to have an association with the search tactic preferred. Ross (2001) found that when choosing books to read for pleasure, readers used a variety of considerations to identify interesting novels. Ross refers to these considerations as "behind the eyes" knowledge where previous experience and meta-knowledge of authors, publishers, cover-art, conventions for promoting books and recommendations from family or friends were involved when considering for selection or rejection any particular book (Ross 2001.) According to Ross, reader's behind the eyes knowledge is involved in a search process in a way that more competent readers are able to rely more on the behind the eyes knowledge gained. Thus, when choosing books the less competent readers rely more on clues in books themselves like the cover design or the blurb (Ross 2001.)

Based on her studies Ross (2001) suggested a model for the process of choosing a book to read for pleasure. According to Ross, the book selection process includes the following elements:

1. *Reading experience wanted*, which reflects the reader's mood when selecting a book

2. *Alerting sources* that the reader uses to find hints of possibly interesting readings, for example recommendations from friends and acquaintances

3. *Elements and clues in the book itself* that readers evaluate when considering whether the book matches the reading experience wanted

4. *Cost in time and money* for the reader in getting an access to a particular book.

In a qualitative study of 16 fiction readers in a public library, Saarinen & Vakkari (2013) found that the readers' typical search tactics varied by reading activity. Avid readers, characterized by a constant reading of various types of novels, were using both known item search and browsing strategies, whereas occasional readers, to whom reading offered occasional relaxation and who read only periodically, were using known item search strategy only. Occasional readers typically knew in advance what they wished to borrow in the library, which is typical of known item search (Saarinen & Vakkari 2013). When selecting novels in public libraries, fiction readers actively searched for the clues of good novels in the books themselves. The authors suggested that as in Goodall (1989), the indicators of good novels could be categorized into clues that trigger interest toward a novel; and clues that determine the decision of borrowing the novel. The most important indicators of good novels

were the author's name, back cover text and a within book browsing, i.e. scanning a particular item.

Ooi & Liew (2011) found when studying fiction selection of 12 adult fiction readers in a public library that participants tended to have a book title or author already in mind when they visited their public library, and fiction selection was thus often based on a pre-determined decision. Instead of browsing in public libraries, participants sought actively recommendations for interesting books outside libraries, where trusted family members and friends, book club discussions, mass media and book reviews played a major role (Ooi & Liew 2011.) These findings are consistent to *alerting sources* and *cost in time and money* as presented in the Ross's (2001) model of fiction readers' book selection. Browsing of fiction in Ooi & Liew (2011) was usually focused on the "New Books" display and the "Book returns" section. Similarly Saarinen & Vakkari (2013) found that the shelf for returned books was a common place for avid readers to begin their search. When searching for fiction by browsing, readers' literary competence plays a major role: less competent readers' natural source of information is the book itself while more competent readers may rely on their knowledge on authors, titles or genres acquired through a constant scanning of their everyday environments (Ross 2001; Saarinen & Vakkari 2013).

Searching for non-fiction in public libraries is characterized by somewhat different factors than fiction searching. Readers of imaginary literature do not typically read fiction books based on the book's topic (Waples 1940), whereas the importance of a subject approach is overwhelming in non-fiction book search with a great majority of users searching for a non-fiction book on a particular subject (Timperley & Spiller 1999). McKay & Conyers (2010) examined library users' difficulties searching the shelves and revealed browsing by subject to be a natural means of obtaining non-fiction books. The authors suggest that subject browsing should be well supported in libraries. Rowlands & Nicholas (2008) conducted a survey on the book selection strategies applied by the members of an academic community to identify the books they needed for work, study, or leisure. The results revealed a range of selection methods from searching a library catalog or Amazon, visiting a library in person to book reviews and personal recommendations. The findings also revealed the modes of book discovery to be highly associated with gender and subject discipline (Rowlands & Nicholas 2008). Mckay et al. (2014) used a circulation data of an academic library consortium to investigate whether books on the library stacks were loaned in clusters. The findings showed a significant neighbor effect in book loans, where a nearby neighbor being loaned greatly increased the probability of any book being loaned itself. The result supports the notion that like

in fiction searching, browsing is a major means to obtain non-fiction books. The finding slightly contradicts the notion that serendipity would have a much smaller role in non-fiction searching compared to fiction searching (Timperley & Spiller 1999).

Summary of book search behavior at the shelves of public libraries. To summarize, previous studies have illuminated two major search approaches for fiction books at the shelves of public libraries: a known item search and browsing. The selection of a particular search strategy has shown to be associated to the frequency of fiction reading, to the reader's literary competence as in previous knowledge on preferred authors and titles, to affective factors as in reading experience wanted and to the elements in the books that trigger interest towards a novel. Next, the search behavior for fiction is examined in digital environments.

2.3   Searching for fiction books in online library catalogs

This section outlines previous empirical studies regarding fiction book search in digital environments. Current knowledge on users' search behavior with digital library systems has mainly focused on retrieving for non-fiction (Lau & Goh 2006; Moulaison 2008; Waller 2010; Ville´n-Rueda et al. 2007). Also in system development the focus has been on retrieving for non-fiction (Case 2007; Elsweiler et al. 2011). Although fiction searching occurs increasingly in digital environments, in modern library catalogs and digital libraries in particular, there is only scarce information on how fiction readers interact with them. The previous section revealed that in selecting fiction in public libraries, readers tend to rely a) to their previous knowledge on preferred authors and titles, b) browsing the shelves for interesting book discoveries, and c) to the clues in the books themselves. In digital environments, the lack of a traditional physical library space where neither the shelves to browse nor the books to skim are available may have an effect on the book selection process. This section focuses on the previous knowledge on search tactics for fiction in library catalogs with a detailed insight into various browsing search tactics. Moreover, the studies on the role of metadata in fiction selecting are described.

2.3.1 Search tactics for fiction in library catalogs

Overall, we have only few studies (Ekvall & Larsson 1997; Jiang 2013; Koolen et al. 2015; Pejtersen & Austin 1983 & 1984; Pejtersen 1989; Oksanen & Vakkari 2012; Pöntinen & Vakkari 2013; Thudt et al 2012; Vakkari et al. 2014) describing fiction book selection within digital library systems. In this study, the focus in examining fiction book search in digital environment is on fiction readers' *search moves* and *search tactics*. Search moves and search tactics in online searching have been examined by (among others) Bates (1979, 1990), Fidel (1985), Wildemuth (2004) and Xie and Joo (2010). Fidel (1985) described a search move as a query formulation during a search process. By analyzing 90 searches, she identified operational and conceptual search moves. Operational search moves modified the query without changing the meaning of query components, whereas conceptual search moves changed the meaning of query components. Bates (1979) identified 29 search tactics in four categories: monitoring, file structure, search formulation, and term tactics. In a subsequent research, Bates (1990) presented levels of search activities from a search move to a search strategy. A search move refers to an identifiable action that is a part of information searching. According to Wildemuth (2004), a search move refers to an iteration in the search formulation and reformulation process. Instead of analyzing individual search moves, Wildemuth examined the sequences of moves. A set of moves that were temporally and semantically related was described as a search tactic. Xie and Joo (2010) analyzed the transitions of search tactics during online searching. They defined search moves as basic actions in the information search process. Search moves formulate search tactics, which refer to users' choices that are applied to advance the search process.

Search tactics for fiction books were studied in detail by Pejtersen (1989) who presented five search tactics applied for fiction searching and a model of a system that was designed to support these tactics, the BOOKHOUSE model. In the BOOKHOUSE model, *a bibliographical search tactic* was used to search for a known item. The BOOKHOUSE aimed at verifying user's information about a known item and locating the item physically. *An analytical search tactic* was used when a user wished access books about a topic such as multiculturalism. It was supported by a bibliographical database including books indexed with the AMP classification scheme which characterized books' attributes with four different dimensions: subject matter, frame, author's intention and accessibility (Pejtersen & Austin 1983 & 1984;

Pejtersen 1989). *Search by analogy* was generated when a user wanted something similar to a novel they had previously read. Similarity was realized via intermediary who discussed with the reader in order to identify the book's specific features which the reader had found interesting. Based on this information, a database was used to find books that shared these features using keywords. A browsing tactic was applied in situations when a user had only a vague idea of what they would have liked to read. *The browsing tactic* was supported via different starting points to collection, such as bibliographical data and the lists of subject terms. *An empirical search tactic* was based on intermediary's empirical experience. An intermediary aimed at recognizing the user's need by a combination of questions and socioeconomical factors and classified users in to characteristic categories with a range of typical set of genres and book titles (Pejtersen & Austin 1983 & 1984, Pejtersen 1989.)

The BOOKHOUSE model was evaluated at a single public library in Denmark in 1988 by real users. It received a positive response with regard to the AMP classification scheme and the icon-based user interface. Users were especially satisfied with the indexing principles used for the concept analysis of documents as they were effectively matching the users' perception of the contents of the books. Also the ability to use different search tactics gave the users feel of system flexibility as it enabled different routes to the contents of the collection. The most popular search tactics used in BOOKHOUSE evaluation were the analytical search tactics and search by browsing the images of book covers (Pejtersen 1989.)

More recently, to support fiction readers in their exploratory quest for good books, various platforms for a fiction book search such as NoveList and LibraryThing have emerged. In NoveList, appealing factors in fiction books such as a storyline, a writing style, a pace and a tone have been added in a catalog to describe the experience of a book for the reader (Caplinger 2013). In LibraryThing, users may catalogue books using data drawn from sources such as the Amazon and the Library of Congress. LibraryThing has been examined by Richards & Sen (2013) and Iyer & Bungo (2011) and it has been shown to be an effective tool for libraries to promote collections (Richards & Sen 2013). The diverse tag categories in Library Thing may support readers in choosing new books as they provide information beyond that of the subject headings (Iyer & Bungo 2011). However, in spite of the emerge on various tools for fiction book search, they have not been evaluated from a user perspective. Thus, there remains to be a gap in our understanding on how these tools support readers to find books to read.

Previous research concerning fiction book selection (i.e. Pejtersen 1989; Spiller 1980; Yu and O'Brien 1996) in public libraries has revealed that two major

approaches to fiction book search are selecting known titles or authors and browsing for possibly interesting items without a clearly defined idea. Adkins and Bossaller (2007) compared different entry points to fiction collections in online bookstores, reader advisory databases and public library catalogs. Known item search was the most supported selection strategy by library catalogs. The authors suggested that public library catalogs can function effectively, especially for known-item searches. In selecting novels by browsing, the process might be guided by a particular topic or a previously read interesting title or it might be untargeted aiming to find just interesting books (Pejtersen 1989; Thudt et al. 2012). Hinze et al. (2012) observed the book selection process in academic libraries and found that browsing and serendipitous discoveries were not effectively supported in digital libraries. The authors point out that browsing in libraries, whether it physical or digital, is an activity that relies strongly on serendipity. The authors suggest means of better supporting serendipitous searching in digital libraries' collections. For example, a one-dimensional list-based display of library collection creates similar difficulties as browsing of physical bookshelves as only limited part of the collection is visible at the same time. The authors suggest that ebooks could be displayed in a two-dimensional space representing similarities between books according to search criteria. The content of the collection could also be displayed in a way that the displayed items on the collection are constantly changing, a way that is called *streaming collage display* by Bainbridge et al. (2004). Streaming collage displays has been demonstrated to be effective for collection understanding and serendipitous discoveries (Cunningham & Bennett 2008; Hinze et al. 2012).

Oksanen and Vakkari (2012) examined fiction readers' search approaches in an enriched public library catalog. They found that issuing queries, exploring result lists, and inspecting book pages were the major search moves for fiction. The findings showed that the effort invested in inspecting search results and book pages instead of querying was an essential factor for finding interesting novels in browsing situations. Also, Vakkari & Pöntinen (2015) evaluated users' search result page (SERP) browsing patterns between an enriched and a traditional library catalog. The results showed that the enriched catalog supported users to identify potentially clickable items on the results list sooner and more effectively compared with a traditional library catalog.

2.3.2   Browsing for fiction in library catalogs

The adoption of web 2.0 technologies in library catalogs has expanded the concept of browsing to include a variety of means to access books without issuing a query as the web 2.0 technologies focus on discovery rather than simply searching for library materials (Emanuel 2011). The common used web 2.0 technologies in library catalogs include user generated tags, automatic book recommendations and visualizations of book collections. It has been suggested that browsing may benefit from visual interfaces that allow for flexible navigation through data collections. Also, search strategies such as exploratory search and information encountering have been recommended as more likely to support serendipity, which is often inherent when searching for fiction (Thudt et al. 2012). A library catalog offering different perspectives to the fiction collection and allowing visual browsing called The Bohemian bookshelf was introduced by Thudt et al. (2012). The Bohemian bookshelf consisted of the images of book covers, keyword chain visualizations, virtual book piles, book recommendations and connected visualizations between books. These attributes simulated a physical library environment, where one can browse the book shelves and the book displays in order to find interesting books via serendipitous discoveries. When designing the Bohemian bookshelf, it was assumed that physical and virtual attributes play a major role during open-ended explorations of book collections. Thus, the design of Bohemian bookshelf aimed at providing a variety of perspectives or facets to the book collection. The Bohemian bookshelf was evaluated by eleven random library users. It was found to support especially open-ended browsing. Users gave the positive feedback of the system's way of presenting books, especially of the multiple visual access points to the collection (Thudt et al. 2012.)

Bawden (2011) distinguishes between active and passive browsing. In active browsing, also referred as targeted browsing (Thudt et al. 2012), the reader has some idea of the book he wants to find, such as the subject of the book. Most search interfaces to digital libraries support targeted search in the form of querying (Marchionini 2006; Rice 1988; Toms 2000). In passive browsing, also referred as open-ended browsing (Thudt et al. 2012), the reader wants to find just good books. The concepts of *social browsing* and *pivot browsing* (Peters 2011; Tang et al. 2014) refer to the ability to re-orientate browsing by social navigational tools and various features realized with web 2.0 tools, such as tags, leading the user to all resources

indexed with certain tags. Tang et al. (2014) conducted a user study of aNobii, an online book sharing website. It provides three social book-finding tools: browsing books by known authors, browsing similar bookshelves, and browsing friends' bookshelves. The author browsing was found to be the most efficient, while browsing similar and friends' bookshelves produced more serendipitous choices. In the same study, users' preference characteristics association with the fiction book selection was also examined. The results suggested that diverse preferences for recreational reading material were associated with different user behavior in a social book finding tool. Readers with high knowledge about their reading interest relied strongly on author browsing while readers with high motivation to seek out appealing readings favored browsing friends' bookshelves (Tang et al. 2014). Furthermore, in a study of the association on positive versus negative online book reviews to fiction readers' book judgments it was found that readers with high motivation for new discoveries were less likely to be influenced by online reviews (Tang 2016). Jiang (2013) examined the users of a social library system and noticed that in addition to the internal search engine, catalog browsing, associative browsing and social browsing were popular search strategies. Searching was found to be the most adopted mode, while browsing appeared to be the most widespread mode.

2.3.3   Role of metadata in fiction selecting in library catalogs

The experience of within book browsing, often crucial while selecting books in public libraries, might be partly substituted by displaying a rich metadata in an online library catalog. Instead of skimming a book in reality, the clues of interesting readings may be searched from the publication information and the content description displayed at the search results and at the item pages. The association of metadata and fiction book selections has been studied by Pöntinen & Vakkari (2013), Vakkari et al. (2014) and Koolen et al. (2015). In their gaze tracking study, Pöntinen and Vakkari (2013) investigated how the metadata of books was examined in an enriched library catalog compared to a traditional one while searching for fiction. They found that the same metadata elements were inspected as much in both catalogs. The strongest predictor of successful book selection was the duration of visits to the publication information, to the author/title information in particular. The finding challenged the importance of the content description (keywords, a text sample, a back cover text, a book review, a cover image) in fiction selection. The significance of the metadata in

fiction book selection was revealed to be associated to cases where the readers had difficulties in deciding the value of a particular book (Vakkari et al. 2014). In deciding a book's interest level between a not interesting, a somewhat interesting and a very interesting, most time for examining the metadata and reflecting on book information was needed in somewhat interesting cases. The metadata displayed already at the search results seems also be associated to successful fiction book selections (Vakkari & Pöntinen 2015). It was found by Vakkari & Pöntinen (2015) that a metadata enriched result list supported readers to identify more efficiently potentially clickable and interesting items compared to a traditional search result list displaying the title, the author and the publication year of a particular item. Errors in a metadata presentation may also have a negative association to a book selection process. McKay et al. (2012) examined the influence on an e-book presentation on the length of reader interactions with e-books. Inconsistencies or errors in a metadata (a cover image, a blurb and a table of contents) were found to be associated to an undesirable short time-span reading: flaws encouraged the user to loan the book for a closer inspection but ended up in rejection rapidly.

In addition to professional metadata, the adoption of web 2.0 technologies and social media in library catalogs allows users and readers to participate in the creating of versatile metadata with tags, book reviews and book recommendations. Koolen et al. (2015) compared the usefulness of professional metadata and user-generated content in determining books' topical relevance in social book search. Findings show that professional metadata was often not enough to determine books' topical relevance. Also, user-generated content was more effective for social book search compared to professional metadata. Weaver (2007) surveyed the tags provided by the readers of a particular novel, The Da Vinci Code. The resulting collection of tags provided a richer description of the book than did a particular social book-related web site. Weaver found that the tag collections provided by users can be divided into different categories, each reflecting a different facet of a novel: character, plot, topic, setting and genre. Weaver (2007) suggested that the use of tags as a part of fiction content description in public library catalogs could allow users to express and meet their needs towards fiction in a wider way than traditional, library professional created metadata would do. However, there remains to be a gap in our understanding about users' perceptions and usability of tags as a part of fiction selection.

Summary of searching for fiction in online library catalogs. To sum up, previous empirical studies have revealed that a known item search is a popular approach to fiction also in an online environment, and is well-supported by current

library catalogs. Browsing as a means to accessing books in digital environments contains various approaches such as social, visual, and pivot browsing. Knowledge on user behavior and search success while browsing for fiction in library catalogs remains to be limited. In addition to a known item search and browsing, topical search for novels and selecting novels based on similar reading experiences have been detected in previous studies. Rich metadata, both professional and user-generated seems to support the fiction book selection in an online environment.

2.4   Searching for books as a casual leisure activity

The aim of this section is to describe the fiction book search as a casual leisure activity and reveal the differences in fiction book search in comparison with a search for non-fiction information resources. According to Elsweiler et al. (2012), *"during casual-leisure search, peoples' intentions, their motivations, their criteria for success, and their querying behaviour all differ from typical web search"*. Previous empirical studies on casual-leisure search have examined information needs in a context of television viewing (Elsweiler et al. 2010), mobile search behavior for leisure events (Schaller et al. 2012), music information seeking in everyday-life context (Laplante & Downie 2011) and the tasks motivating twitter search (Elsweiler & Harvey 2015). Previous empirical studies on fiction book search as a casual-leisure activity remain unknown. However, searching for fiction books may be considered as a casual leisure activity that is not motivated by an actual information need or a task to be solved. Instead, personal, pleasurable and hedonistic objectives engage fiction readers in often exploratory searching behavior that combines querying and browsing strategies for investigating interesting books to read (Elsweiler et al. 2011; Marchionini 2006).

It is particularly the absence of a specific information need that distinguishes the fiction book search from a task-oriented or a work related information behavior. The information need, the obvious goal that launches the search process for recreational reading material might simply be "to find good books" or find books that entertain, provide relaxation and escape from daily routines. As both the intention and the purpose of the search are often ill-defined, it is challenging to evaluate what counts as a successful searching session. During a casual-leisure search,

the journey and the experience while searching are often important in itself (Elsweiler et al. 2012.)

According to Elsweiler et al. (2012), the importance of traditional information retrieval algorithms and the appropriateness of common evaluation metrics such as time, precision and recall in the casual-leisure search should be questioned. When users are spending time and enjoying themselves while navigating through a digital library in a quest for interesting readings, should it create a misinterpretation if more time spent was considered as a flaw in evaluating the search process? Similarly, the question of a document's relevance in the casual-leisure search could be challenged. The concept of relevance has been acknowledged as central to the theory of information retrieval and a fundamental concern in evaluating information retrieval systems (Borlund 2003; Cooper 1971). The studies on relevance have focused either on a system-centered perspective on relevance as a logical and topical relationship between a user's query and a subject of a document, or on a user-centered approach (Borlund 2003; Cooper 1971; Saracevic 1996). Saracevic (1996) distinguishes between five basic types of relevance: algorithmic, topical, cognitive, situational and affective relevance. Saracevic's affective relevance as the relation between the intents, goals and motivations of the user and the document fits to the emotion-bound fiction selection process (Ross 2001). However, affective relevance in fiction book selection has not been examined.

Previous research on relevance assessments in fiction book selection in library catalogs has been conducted by Reuter (2007) and Koolen et al. (2015). Reuter (2007) examined children's book selection in a digital library with the concept of "aesthetic relevance." Aesthetic relevance was defined as the potential of a document to provide a suitable reading experience. The most influential factors in children's book selections were the metadata (such as title, author), the extrinsic appeal of a book (such as recency, format) and accessibility (such as length, text density). Koolen et al. (2015) investigated the relevance aspects expressed in book requests in the Library Thing discussion forums. As in Reuter (2007), the authors identified relevance aspects such as accessibility, content, engagement, familiarity, metadata, novelty, known item and sociocultural. The most adopted relevance aspects in book requests were the content of a book and looking for familiar reading experiences. Also, a unique combination of content, context and examples in book requests was detected.

Summary of fiction searching as a casual leisure activity. Overall, the fiction book search as a casual-leisure activity remains to be an under-researched area. The central differences in the fiction book search in comparison with a search for non-fiction information resources are related to

the absence of a specific information need,

the affective factors while selecting books,

the challenge in evaluating the search success and the relevance of the possible book discoveries.

The information need in a fiction book search may not be specific. Yet the natural goal of the search is the finding of an interesting item. Even though the journey through an online environment for fiction books would be experienced as amusing or satisfying by the users, if it does not lead to any book discoveries, the search is considered as unsuccessful in this study. Here I transfer the aspect of evaluating the relevance of the items found to the readers themselves: when selecting fiction books in this study, the readers will evaluate the interest level of selected items from their own perspective.

3 THE RESEARCH DESIGN

This chapter describes the empirical work that was carried out. The dissertation consists of two separate data sets:

- a survey data collected by Vakkari & Serola in the year 2010 for studying the outcomes of public libraries analyzed in Study I, and

- a user-study conducted in the year 2012 for studying the search behaviors of fiction readers in online environments analyzed in Studies II-V.

The first section describes the participants and the data collection for Study I. The second and third sections illustrate the recruitment of participants, the library catalogs used and the experimental procedure in Studies II-V. Section 3.4 explains how the data collected were processed and analyzed. The chapter ends with discussion of the validity, reliability, and ethics of the research.

3.1   The participants and data collection in Study I

The primary motivation of the Study I was to survey the contributions that libraries as social institutions make to society at large (see Vakkari & Serola 2012). Hence, a survey questionnaire measuring the outcomes of public libraries was designed and directed to the Finnish adult population. Alternative objectives for the data collection, such as interviews, were also considered. However, a survey questionnaire was considered as the most suitable method for answering the research problems.

The data was collected in 2010 as a stratified random sample of the Finnish population aged 15 to 79 years by mail questionnaire for studying the outcomes of public libraries (Vakkari & Serola 2012). The number of respondents was 1000. Compared to the overall Finnish population aged 15 to 79 years, the sample was well represented by age, marital status, and geographic region (Statistics Finland 2009) but it was biased toward female and highly educated (see Vakkari & Serola 2012). Studies (Huysmans & Hillebrink 2008; Ross et al. 2006; Vakkari 1991) indicate that females are more frequent library users compared to males, and highly educated

people are more frequent library users compared to less educated people. By implication, the sample was biased towards more active library users. The greater response rate of those interested in the phenomenon observed is a common feature in surveys (Groves 2006). More information of the sample, see Vakkari & Serola (2012).

The survey included questions concerning search tactics for accessing books in public libraries, respondents' reading activity and the type of literature read in the prior 12 months. The questionnaire contained also variables measuring demographic factors age, gender and educational level. These variables were chosen in the study as it has been shown that age, gender and educational level are associated to pleasure readers' reading activity and reading preferences (Griswold et al. 2005; Jennings & Sear 1986; Moyer 2007; Ross 2001; Ross et al. 2006; Spiller 1980; Statistics Finland 2009; Tepper 2000).

The major search tactics in public libraries were derived from the previous literature (Goodall 1989; Pejtersen 1989; Spiller 1980; Yu & O'Brien 1996) which were known title or author search, browsing on the shelves, browsing on the shelves of returned books, catalogue use, asking librarians and browsing book displays. These tactics were represented as follows in the questionnaire: *I know the name of the book or the author, I browse on the shelves, I skim the returned loans, I use the library catalog, I ask a librarian and I browse book displays.*

The respondents were asked to indicate how often they used each of the tactics for searching for books in public libraries in an ordinal scale ranging from 1 to 4, where 4 was "often", 3 was "sometimes", 2 was "seldom" and 1 was "never". An option "I can't tell" was also given. The measure used indicates search tactics used for finding books in general in public libraries. Thus the respondents were not asked how often they used these tactics for finding solely fiction. However, as the majority of books borrowed in public libraries are fiction, the measure reflects mostly tactics for finding fiction. Moreover, as these six search tactics were derived from the previous literature with focus on fiction searching, it could be assumed that they reflect tactics applied for fiction searching.

As an indication of reading activity the respondents were asked to mark the number of books read in the last 12 months, excluding textbooks read for studying. The respondents were also asked to indicate how many of these books were non-fiction books. With this information new variables, the number of fiction books read in the last 12 months and the percentage of fiction books of the total amount of books read in the last 12 months, were constructed.

Demographic variables included in the Study I were gender, age and the level of education. For more information about these variables see Vakkari & Serola (2012). The educational level of the respondents was divided into three categories, referred to here as low, middle, and high. Low is primary school education which consists of eight years of studies. As the previous comprehensive school education in Finland prior to school reform in the 1970s, only the older population has studied under this system. Middle level education is comprehensive school education, which is general knowledge education. This compulsory education starts in the year of the child's seventh birthday and is finished when the syllabus of comprehensive school education has been completed, 9 years later. High education refers to education consisting of at least Matriculation examination and international matriculation examinations (IB, Gymnasieexamen), which is three to four years of studies after the comprehensive school.

3.2 The participants in Studies II-V

The data collection for the Studies II-V took place in 2012. Participants were recruited in three cities situated in the Pirkanmaa, Kanta-Häme, and Uusimaa regions in Finland. In order to find participants with real fiction reading interest, 80 people were recruited in public libraries, in fiction reading groups, and in writing and literature classes in the Open University of Finland. The writing and literature classes were open to all adult students paying a course fee, regardless of age or educational background. The literature and writing classes were arranged in cooperation with university departments. In addition, the snowball sampling method and a newspaper advertisement were used. Participants were offered a movie voucher to participate in the study. 80 participants were considered a group large enough to conduct various between group analysis. If a participant had not read a single fiction book during previous 12 months and reported personally disinterest in fiction reading, s/he were excluded from the study due to lack of genuine fiction reading interest.

The participants were randomized into control and test groups, where the control group used a traditional online catalog and the test group used an enhanced catalog that provided enriched fiction related metadata elements (such as cover images and book descriptions) and enabled users to browse the collection via different access points to literature.

In both groups, 18% of the participants were male and 82% female. In both groups, the age distribution of participants varied from 20 to 80 (Figure 2). In the group using a traditional catalog, participants averaged 34 years of age (SD 12.7). In the case of the enriched catalog, participants averaged 42 years of age (SD 16.8). The age distribution of participants in the enriched catalog group was significantly higher in comparison with the age distribution in the traditional catalog group ($t = 2.46$, $p < 0.05$). However, as there was no significant difference in participants' experience in using online catalogs between control and test groups, the difference unlikely affected the results.
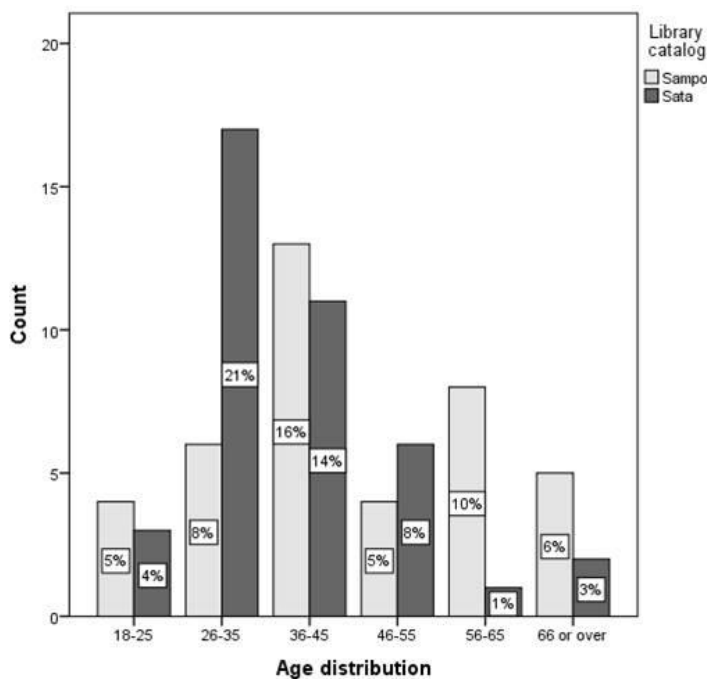
Figure 2. Age distribution of participants.

The educational level of the participants was designated as middle or high (see page 31). In both groups, 18% of the participants had a middle level education and 82% a high level education. 81% of the participants in the traditional catalog group and 58% of the participants in the enriched catalog group used online catalogs for fiction searching at least once every three months. In the traditional catalog group, 5% of the participants had never used online catalogs for fiction searching. In the

enriched catalog group, 10% of the participants had never used online catalogs for fiction searching. No significant differences were found in the use of online catalogs between the groups.

Fiction reading activity and fiction reading preferences did not significantly differ between the control and test group. On average, the participants in the enriched catalog group read 26 novels a year while the respective figure in the traditional catalog group was 18 (t=1.409, p< 0.05). Participants were asked to fill in a pre-questionnaire measuring their reading preferences (Appendix II). The questionnaire was designed to measure motives for fiction reading and important features in fiction texts while reading fiction books.

3.3   Data collection in Studies II-V

This section illustrates the data collection for Studies II-V. It describes the simulated search scenarios, the library catalogs and the experimental procedure applied during the data collection.

3.3.1   The search scenarios

The data collection for the studies II-V was thoroughly considered. Various alternative options for the data collection, such as analyzing real log files, applying the thinking aloud method, or observing participants in natural settings in authentic search situations, were contemplated. The primary aim of the Studies II-V was to compare readers' search behaviors between various catalog and search types. Hence, it was thought as a necessity to maintain a researcher's control over the data collection (in terms of catalogs used and search scenarios applied) in order to enable a reliable comparison between catalogs and search scenarios. However, a possibility for the participants to proceed according to their genuine reading interest in most of the designed search scenarios was also ensured. In addition, a mixture of various data collection methods (analyzing the search logs captured from the screen actions, using various questionnaires, observing and interviewing the participants) was

adopted while collecting the data for creating a comprehensive understanding about the fiction book search process.

For designing the search scenarios for the Studies II-V, previous literature (Adkins & Bossaller 2007; Goodall 1989; Moyer 2007; Peters 2011; Pejtersen 1989; Ross 2001; Spiller 1980; Thudt et al. 2012; Van Riel 1999; Yu & O'Brien 1996) concerning fiction searching in public libraries was thoroughly examined with an attempt to identify the typical contexts where fiction searching in public libraries has commonly occurred. A known item search, a topical search, an open ended search, a search for similar books, a search for a vacation /a trip, a social book search (searching for books from the "Book returns"), a visual book search (searching for books based on cover images from book exhibitions) and a search with a help from a librarian were identified as the typical fiction book search situations occurring in public libraries. When designing the search scenarios for the Studies II-V, each of the identified search situations was considered as a potential for being included in the experiment. However, asking help from a librarian was excluded from the experiment, as the idea was to examine how *library catalogs* support fiction readers in their fiction book search. Visual and social book searches were considered as specific types of open ended browsing, which was included in the experiment of this study as a simulated search scenario.

As a result, five simulated search scenarios were designed. In two search scenarios, the participants were given an author or a topic with which to begin the search process. The remaining three search scenarios explored the concept of individual and dynamic information needs, as the participants were asked to proceed according to their personal preferences without any given topic. Search scenarios reflect the following typical search tactics in fiction searching: known author/title search (Pejtersen 1989; Spiller 1980; Yu & O'Brien 1996), open ended browsing (Bawden 2011; Pejtersen 1989; Thudt et al. 2012), topical search (Bawden 2011; Pejtersen 1989; Thudt et al. 2012), search by analogy (Pejtersen 1989), and searching without conducting a query (Thudt et al. 2012). Simulated search scenarios were as follows:

The known author search: "A friend of yours recommends you to familiarize yourself with the novels of Olli Jalonen. Find Olli Jalonen's novels and choose two novels which are of interest to you."

The topical search: "Find three novels about upper class life in the 19th century."

The open-ended browsing: "Find three novels that interest you which you would like to read."

The search by analogy: "Think of and mention one novel that you have read and found interesting recently. Now search for three novels that you would consider similarly interesting as the one you mentioned."

The open-ended browsing for vacation reading: "You are leaving for a holiday trip. Find three entertaining novels for the trip without using the free text box."

In the scenario of open-ended browsing for vacation reading, the search began without conducting a query. In the test group, the participants were instructed to begin the search by using the catalog's other features instead of issuing a query. For the control group using the traditional system, the scenarios was presented in a slightly different way, as it was not possible to browse the collection without first conducting a query. The statement for the fifth scenario in the control system was: "You are leaving for a holiday trip. Find three entertaining novels for the trip."

Depending on the precise research problems presented in each of the four studies (II-V), solely the search scenarios where participants' search behavior answered the research questions were included in the data analysis. In Study II, each of the search scenarios was included in the data analysis. In Studies III and IV, three of the search scenarios and in Study V four of the search scenarios were analyzed.

3.3.2   The library catalogs

Satakirjastot (Sata) is the web service of the city libraries of the Satakunta region in Finland. Since it includes a traditional public library online catalog, it was used as the baseline system. The service consists of a library catalog, information retrieval system, customer record, basic and advanced search from a variety of databases, and links to each library's own web services. In Sata, the metadata for fiction includes bibliographic information and subject terms from the fiction thesaurus Kaunokki (Saarti & Hypén 2010). For books published recently, cover images and blurbs are also available. The search options include basic and advanced search. Basic search matches the given words in every accessible part of the data. In advanced search, it is possible to narrow the search by several options, including specific metadata fields, in the dropdown menu (Figure 3).

Figure 3. Advanced search in Satakirjastot.

The search results are sorted by format (such as book, DVD, CD) and relevance. After narrowing down to the preferred format, the search results can be sorted by author, title, publication year, or class. Figure 4 presents the search result page in Sata.

Figure 4. Search result page in Satakirjastot.

BookSampo-service (Sampo) was used as the enhanced catalog (the test group) for the study. Compared to the baseline catalog Satakirjastot, Sampo provides fiction

related metadata that is more varied and allows for different access points to literature (Table 1). Sampo contains metadata for the adult fiction collection from the HelMetWeb Library, which is the web service of the city libraries of the Helsinki Metropolitan area in Finland. In Sampo, the associations and similarities between the works of literature are realized by semantic web technologies such as the ontologization of the fiction thesaurus Kaunokki, which is used for fiction indexing in Sampo (Hypén & Mäkelä 2011).

Table 1. The major characteristics of the catalogs used.

Characteristics		Sata	Sampo
Searching	Basic search	Yes	Yes
	Advanced search	Yes	No
	Cover image search	No	Yes
	Browsing	Yes, after querying	Yes
Book page information	Author, title	Yes, always	Yes, always
	Keywords	Yes, often	Yes, often
	Blurb	Yes, seldom	Yes, often
	Cover image	Yes, seldom	Yes, often

The front page of Sampo offers a variety of access points to the collection. Users can begin the search process with a basic search (similar to the baseline catalog, Figure 5), by browsing cover images in a book cover carousel, by skimming other users' virtual bookshelves, or by looking at a subject term cloud (Figure 6).

Legend: 1=Basic search, 2= Results by author, 3= Results by genre, 4=Snippet

Figure 5. Figure 5. Basic search and search result page in BookSampo.

Figure 6. BookSampo book cover carousel (1) and subject term cloud (2).

The book cover carousel (Figure 6) on the front page displays 25 randomly picked items which change daily. Further, five randomly picked virtual bookshelves created by other users are displayed on the front page, with an option to view all the user-

created bookshelves. Book reviews and book recommendations are provided daily by the administrators of the service, also on the front page.

The search results are sorted in two separate sorts: one by author and another by genre of literature (novels, plays, short stories). A book page represents the content of a particular novel in detail: metadata on subject and themes, milieus, settings, main characters, and time frames are included alongside traditional subject terms and cover images. Book descriptions, automatic recommendations, and text samples are also included.

When the catalogs for the experiment were considered, the aim was to identify a traditional catalog with a fiction collection comparable to Sampo in size and content. The collection in Sata was found to be relatively similar to Sampo's collection. The fiction collection of Sata consisted of 77,000 titles and the fiction collection of Sampo of 73,000 titles, both including novels, plays, and poems. Both collections included the classics and recently published titles of Finnish and foreign literature.

Before the experiment, the researchers conducted several test queries in both catalogs in order to compare the results. They were found to share several overlapping titles. When the data was collected, e-books were not available in the fiction collections of Sata and Sampo. The analysis of Sata and Sampo revealed that the size and content of these collections were comparable, and therefore they could be used in the experiment.

3.3.3   Experimental procedure

The experimental setting shared similarities with the work of Vakkari & Pöntinen (2015) and Vakkari et al. (2014). In these studies, the instructions and post-task questions for fiction search tasks were presented in a similar fashion as in this study. Information of the above-mentioned experiment was used when designing the experimental setting of this study. The experimental setting was pre-tested with one participant to gain information on the duration of the test as a whole and to see if the instructions were unambiguous enough.

The search logs were saved with a Morae Recorder (http://www.techsmith.com/morae.html). The Morae Recorder captures audio, video and on-screen activity during a research session. Variables measuring the search actions were manually calculated from on-screen activity after conducting the user test. The audio was used in analyzing and classifying participants' search queries.

The experiment consisted of the following steps:

1. Pre-questionnaire including demographic questions, questions concerning participants' search experience in online catalogs in general, participants' familiarity with the catalog used in the experiment, and their reading interest (Appendixes I and II).

2. Introduction to the experiment.

3. Brief demo of the retrieval system (approximately two minutes long).

4. Execution of five search scenarios.

5. A post-task questionnaire after each completed search scenario.

6. A post-session questionnaire after the completion of all search scenarios.

7. Brief post-task interview after one search scenario (decided in advance).

Based on the findings by Miesen (2003) and Saarinen & Vakkari (2013), a pre-questionnaire measuring participants' reading preferences and motives for fiction reading was designed (Appendix II).

The time for completing the tasks was not limited. As mentioned above, the participants were randomized into control and test groups. Each participant completed the tasks either on Sampo or Sata. The pre-questionnaire confirmed that the participants in both groups were unfamiliar with the catalogs used in the experiment. The known author task functioned as a training task and was conducted first with each participant. Latin square rotation was used with the four other tasks. Latin square technique ensures that each task appears at each position equally often and that each task precedes and follows every other task equally often (Carter & Lubinsky 2016).

Each participant completed the tasks individually. Public libraries, meeting rooms at a university, coffee shops and participants' homes were chosen as test environments according to participants' wishes. During the experiment, the researcher was present to help in case technical problems occurred. At the beginning of the search process, participants were not asked to follow the think-aloud strategy. As the experimental procedure contained various questionnaires and complex search scenarios, it was considered as too stressful for the participants to say out loud what they were thinking. However, during the search process, a majority of the participants began naturally to "think-aloud" and give comments about their decisions while completing the search scenarios. The researcher also asked questions on particular book choices during the search process. However, the researcher did not guide or help the participants in questions concerning the completion of the

search scenarios. In each scenario, the participants were asked to search for three novels that were of interest to them except in the known author scenario, in which the participants were asked to find only two interesting novels. It was thought that finding two interesting novels in the training scenario was enough to familiarize the participants with the catalog's basic functions.

When designing the experiment, the suitability of the concept of (topical) relevance for measuring fiction readers' search success was considered. According to Saracevic (2007), relevance functions as the fundamental notion in information science, particularly in information retrieval (IR). In the traditional view, both the IR systems and their users create relevance. IR systems take a query, match it to information objects in the system and provide what they consider relevant. After that, users derive relevance from the obtained information objects according to their current search task (Saracevic 2007). In the fiction book search, emotions and serendipity play a major role in the search process and readers' search behavior is often untargeted as their interaction with an information system occurs without a topical search objective (Agosti et al. 2014; Ross 2001). Also, fiction readers do not commonly select fiction books based on their topics (Waples 1940). Relevance "as a relation between the subject of the document and the subject of the query" (Saracevic 2007) or relevance as "whether a piece of information is on a subject which has some topical bearing on the information need" (Cooper 1971) seemed ill-suited for the selection of recreational reading materials. Then again, interest as a feeling that accompanies a special attention to an object, or interest as something that draws attention and causes the curiosity of a person, seemed better-suited for the fiction book selection in comparison with (topical) relevance. Interest does not involve a book to have any topical bearing on the information need, which is often undefined in fiction search. Hence, in this study, interest towards a book was defined as a book's capacity to cause the curiosity to explore it more and to provoke a positive idea of the reading experience possibly provided. It was applied in examining fiction book selections at the present study.

After each search scenario, the participants were asked to rank the novels found according to how much they were of interest to them using an ordinal scale ranging from 1 to 3, where 1 was "a little interesting," 2 was "somewhat interesting," and 3 was "very interesting." A score of 0 was used if an interesting novel was not found.

3.4   Data analysis

Separate analyses were carried out for each of the subsidiary studies, I–V. The general principles of the analysis, applied in each of the studies, are presented next.

3.4.1   Analysis of the questionnaire data in Study I

A major aim of the Study I was to explore how common are the search tactics used for fiction searching in public libraries. The data were analyzed by using one-way analyses of variance (ANOVA) to determine whether age, gender, education and the number and the proportion of fiction books read was related to search tactics used. For the analysis, the proportion of those respondents who had used a particular tactic at least sometimes was calculated. Only those respondents who had read at least one book in the previous 12 months were included in the analysis. This was done as it is not likely that those not reading books would search for books in a library.

In order to analyze more in detail associations between search tactics and reading activity, linear regression analysis was applied. For the economy of analysis, the six search tactics observed were reduced by factor analysis into two major *search strategies*. Factor analysis is a statistical data-reduction technique used to explain variability among observed variables in terms of fewer unobserved variables called factors. The loading indicates the proportion of variance a factor is able to explain of the total variance of a variable. Loadings vary between +1 and -1. The closer the loading is to one, the higher the dependence of the particular variable is from the factor (Tabachnick & Fidell 2013). The eigenvalue indicates the degree to which a factor is able to explain the total variance of the variables included in the analysis. If a factor has a low eigenvalue, then it is contributing little to the explanation of variances in the variables and may be ignored as redundant with more important factors (Tabachnick & Fidell 2013). The lower default value in SPSS for forming a factor is 1. The number of factors can be selected for interpretation either mechanically based on eigenvalue, or theoretically based on a theory or theoretically justified interpretation of the factors.

Linear regression analysis was used to analyze and model the two strategies produced by the factor analysis and independent variables. The latter variables were as follows:

the number of fiction books read in the last 12 months

the number of non-fiction books read in the last 12 months

the percentage of fiction books of all books read in the last 12 months

the frequency of public library visits

respondents' gender, age and educational level.

When constructing regression models, the method Enter was used to add independent variables into the models. Enter is a method where the researcher controls the entry of variables and the advancement of the regression process. In the Enter method the variables within the model are either predicted by research hypotheses or selected based on previous literature (Tabachnick & Fidell 2013). The variables were chosen based on literature in an attempt to produce a model which would yield information to understand which among the independent variables are related to the dependent variable.

3.4.2   Analysis of the reading preference questionnaire data in Study IV

In Study IV, a questionnaire data was analyzed to divide the participants in categories sharing similar reading interests. These reader groups were created based on the reading preferences questionnaire (Appendix II). To identify groups of participants sharing similar responses on the questionnaire, and to distinguish different preference patterns in fiction reading, a cluster analysis was conducted. The aim of cluster analysis is to identify homogenous groups based on their shared characteristics. The clustering was guided by the procedures outlined by Hair et al. (1998). First, the cases with missing data were excluded from the analysis (2 cases). Since the variables used in the cluster analysis did not vary in range, standardizing prior to clustering was not needed. Then, a k-means cluster analysis was conducted.

A k-means cluster analysis is a procedure that is applied when the number of clusters is predefined (Hair et al. 1998). Considering the n-number and findings from

previous research (Miesen 2003, Saarinen & Vakkari 2013, Yu & O'Brien 1999), two clusters was hypothesized as a rational solution. First, each variable in the preference questionnaire was included in the analysis. Then, the variables that did not significantly (p >.05 in the variance analysis) contribute to grouping the participants to clusters were excluded from the analysis. The final two-cluster solution contained 13 variables and seemed to be a good fit to the data.

3.4.3   Analysis of the search log data

The search logs were used in analyzing fiction readers' search behaviors in Studies II, III and IV. The search logs were saved with Morae-software. Table 2 presents the coding schemes for variables measuring fiction readers' search behavior in Studies II, III and IV. In each of these subsidiary studies, success in search scenarios was measured by the books' interest grading given by participants. Average book scores per task were calculated.

In Study II, the search moves applied in fiction searching in five search scenarios were analyzed (Table 2, Study II). A search move refers to an identifiable basic action during the search process (similar to work done by Bates 1990; Xie & Joo 2010). Firstly, the search moves were identified by examining the recorded user tests. In Sampo, the label *pivot browsing* (Peters 2011) was applied to search moves that re-orientated browsing to follow features such as virtual bookshelves, a tag cloud, a cover image carousel, and automatic recommendations. In Sata, pivot browsing was not possible. After identifying the individual search moves, a coding scheme was developed to measure the search moves (Table 2, Study II). The scheme was slightly different between the two catalogs used as they offer different search features. The identified search moves were measured by their frequency and duration. Each of the search scenarios was coded individually. In addition, the measures of the search session were also coded.

Table 2. Coding schemes in Studies II, III and IV.

Coding scheme		The variables
Study II	Search moves in Sata	Conducting queries in basic searching Conducting queries in advanced searching Conducting queries in browsing searching Examining the search result list Exploring an individual book page Reading description of a novel (in a separate window) Examining book cover of a novel (in a separate window)
	Search moves in Sampo	Conducting queries in basic searching Conducting queries within book cover images Examining the search result list Exploring an individual book page Exploring the catalog's front page Pivot browsing (i.e. automatic recommendations, cover images)
	Duration-variables	Time to complete a scenario (Sata, Sampo) Time to submit queries (Sata, Sampo) Time to formulate the first query (Sata, Sampo) Time to formulate the following queries (Sata, Sampo) Time to view the search results (Sata, Sampo) Time to view the book pages (Sata, Sampo) Time to view the cover images (Sata, Sampo) Time to read the description of the book (Sata) Time to view the catalog's front page (Sampo) Time to pivot browsing actions (Sampo)
	Frequency-variables	The number of opened book pages The number of queries The number of individual search moves
Study III	Search actions during QRIs	Duration of QRI (Sata, Sampo) Query time (Sata, Sampo) Total SERP time (Sata, Sampo) Dwell time on book pages (Sata, Sampo) Number of SERP visits (Sata, Sampo) Number of opened book pages (Sata, Sampo) Pivot browsing time (Sampo) Number of pivot browsing actions (Sampo)
Study IV	Measures for search actions	Number of search actions Number of search queries Number of visits to SERPs Number of opened book pages Time per a query Time per a SERP visit Time per a book page

The variables in the coding scheme for Study II (Table 2) were not normally distributed. Thus, Friedman's and Mann-Whitney's non-parametric tests were used to test differences in search moves between scenarios and catalogs.

Study III examined fiction readers' search actions during various query reformulation intervals (QRIs). QRI refers to users' search actions that take place between two queries (Liu et al. 2010). According to Liu et al. (2010) issuing a query represents a searcher's cognitive decision to modify the current search tactic, and is likely to affect the searcher's retrieval actions. The search actions between two queries form units, which can be used e.g. for eliciting information about users' document viewing behavior. In Study III, the aim was to understand how readers' search actions differed between successful and unsuccessful QRIs and which search actions predicted the selecting of very interesting novels compared to less interesting ones.

In Study III, 80 user tests were manually coded in order to identify different QRI types, and measure the search actions during the identified QRI types. In both catalogs three types of query reformulation intervals were identified: a) Successful QRI with very interesting novels selected (named as Successful QRI), b) Successful QRI with a little and somewhat interesting novels selected (named as Somewhat successful QRI) and c) Unsuccessful QRI with no interesting novels selected (named as Unsuccessful QRI). After identifying the QRI types, a coding scheme was developed for search actions. Each possible search action was identified in both catalogs by a single coder. The coding scheme was slightly different between the two catalogs used as they offer different search features (Table 2).

In Sata, options for submitting queries were basic and advanced search. In Sampo, solely the basic search option was provided. In Sampo, browsing the collection was possible without querying. Like in Study II, the concept of *pivot browsing* (Peters 2011) was applied to measure search actions that re-orientated browsing to follow features such as virtual bookshelves, cover images and automatic recommendations. In Sata, pivot browsing was not possible. The search actions were measured by their frequency and duration. Each of the search scenarios was coded individually.

The variables in the coding scheme for Study III (Table 2) were not normally distributed. Thus, Friedman's and Mann-Whitney's non-parametric tests were used to test differences in search actions between QRI types within a task, between tasks and between catalogs.

The search actions in the open ended browsing and search by analogy scenarios did not statistically differ from each other. Thus, for the economy of analysis, they were collapsed into one search scenario referred to here as *a browsing task*. To analyze whether the search actions during QRIs were associated to books' interest grading, multinomial logistic regression analysis was applied. The aim was to identify the search actions that predicted the selection of very interesting novels compared to less interesting ones. Before conducting regression analysis, the data was re-coded using QRI as the unit of observation. Books' interest grading given by participants was used as a dependent variable in multinomial logistic regression analysis. The variables presented at Table 2 were used as independent variables. When constructing the regression models the method Enter was used to add independent variables into the models.

In Study IV, where the focus was in comparing fiction readers' search actions and search success between two different fiction reader groups, the identified search actions were measured by their frequency and duration (Table 2). Independent samples t-test was used to test differences in search actions between reader types in each search scenario. For the economy of the analysis, the number of the search scenarios applied in the analysis in Study IV was reduced to three. Known author scenario, open ended browsing and search by analogy yielded information enough to answer to the research problems presented in Study IV.

3.4.4 Analysis of the audio data and the search queries

In Study V, the focus was on investigating fiction readers' interpretations of the influential factors in fiction book selection in different search scenarios. For this purpose, 80 participants' interviews and conversations with the researcher during the experiment were recorded with Morae-software (www. techsmith.com/morae.html). The audio material was transcribed word for word and qualitatively analyzed. The purpose of the qualitative data analysis was to identify the categories for the interest criteria for novels mentioned by the participants, and apply these criteria for examining participants' book selection behavior. To guide the qualitative data analysis, findings from previous research (Reuter 2007; Ross 2001; Saarinen and

Vakkari 2013) were taken into consideration. Particularly, the coding scheme for children's' book selection in a digital library developed and applied by Reuter (2007) supported the design of an initial coding scheme for this study.

Based on a preliminary review of the transcripts, a coding scheme representing the factors that influenced fiction readers' book selection at each search scenario was developed. The authors discussed and refined the code labels and definitions several times in order to ensure the coding scheme to be unambiguous with no overlapping sub-categories. A protocol for marking the utterances was then developed. An utterance was defined as a sentence or a unit of conversation covering a single aspect of a book selection. Utterance units that covered multiple topics were divided and each unit was given an appropriate code. Each utterance was given a single code, no multiple codes for a single utterance were allowed. The transcripts were then reviewed line by line and utterances were encoded according to the coding scheme. The final coding scheme with code definitions and example utterances is included in Appendix III.

The total number of selection mentions in each dimension was first calculated over search scenarios and catalogs. In addition, the total mentions concerning rejecting a book were calculated for each dimension. Then, the mentions in each dimension were calculated by scenarios and by catalogs. To test for significant differences in the distribution of selection mentions between catalogs, $\chi2$ p-values were calculated.

Search queries. In Study IV, the search actions and search queries of diverse fiction readers were examined and compared between two reader groups, the entertained and the aesthetic. In order to analyze the search queries, the query terms were manually identified and saved in the search logs. The number of terms in each query was calculated and the query was classified either as a successful one or as an unsuccessful one. A query was considered as successful if it resulted in a selection of a novel. A query was considered as unsuccessful if it resulted in no selections. Next, the type of each query was qualitatively analyzed and classified in one of the following categories: author, title, genre, topic, person, place or time frame. Each query was placed in one category only. The categories were taken from the Finnish fiction thesaurus Kaunokki (http://onki.fi/en/browser/overview/kauno) which was used in fiction indexing in both catalogs used in the experiment.

The classification of search queries was conducted by a single coder. Intracoder reliability, which refers to the consistent manner by which the researcher codes, was considered during the coding process. In order to test the ability of the coding

protocol to result in a consistent categorization of content, the classification of search queries was conducted two times by the same coder. During the coding, it was assessed whether any incoherence had occurred in the single coder's understanding or application of the protocol definitions. In addition, data sessions with fellow researchers were held during the coding process. In these sessions, the coding protocol was assessed and discussed with other researchers.

The type of search queries was typically unambiguous as the majority of the queries contained query terms from one query category only. In cases where contradictions occurred, the query was placed in a category that was interpreted as representing a searcher's main target by listening to the audio data.

The search queries were statistically analyzed in each scenario and in both reader groups. The proportion of queries in each category was separately calculated for successful and unsuccessful queries. χ^2-tests were used to test for significant differences in query type between the reader groups.

3.4.5   The validity, reliability, and ethics of the research

This thesis consists of five subsidiary studies that all give a different angle to the fiction book search behavior in online environment. To strengthen the validity of the research, methodological triangulation was applied in the research design. According to Denzin (2006), triangulation can be defined as *"the combination of methodologies in the study of the same phenomenon"* and it is an attempt to strengthen a study by using several kinds of methods or data, including both qualitative and quantitative approaches (Patton 2002). In this thesis, triangulation was applied for explaining the richness and complexity of fiction readers' book selection behavior in online environment, a phenomenon studied scarcely in previous research. The data and methods applied in this study include

the qualitative data (interviews, observation of participants, search queries)

the quantitative data (questionnaires, search logs)

qualitative data analysis (manual coding of the search logs, the analysis of the transcribed audio data with content analysis, the qualitative categorization of the search queries)

quantitative data analysis (measures for search actions analyzed with various statistical analysis methods).

The validity and reliability issues inherent in use of both qualitative and quantitative methods were taken in to consideration in each of the subsidiary studies I-V. In Study I, various statistical data analysis methods were applied to survey the major search strategies in public libraries. The number of respondents in the questionnaire data was considerably high as 1000 respondents fulfilled the questionnaire. 1000 respondents were considered as a group large enough to reliably conduct various multivariate statistical data analysis. Compared to the overall Finnish population aged 15 to 79 years, the sample was well represented by age, marital status, and geographic region but it was biased towards female, highly educated and active library users. Thus, the generalizability of the findings should be acknowledged. In addition, the original questionnaire was designed to study the outcomes of public libraries (Vakkari & Serola 2012) where the questions for search tactics in public libraries constituted one sub-field. The options for search tactics were pre-given to the participants, and the importance of each search tactic was evaluated by the participants. Given an open-ended question for participants to describe their search behavior for books, could have affected the findings.

In the experiment for Studies II-V, simulated search scenarios were designed to study the search behaviors for fiction books in online environments. The principles of interactive information retrieval (IIR) evaluation model (Borlund 2003) were considered in designing the experiment. The key elements of the IIR evaluation model are the use of natural and realistic search scenarios (simulated work task situations) and the alternative performance measures (Borlund 2003). A simulated work task situation is a short "cover story" that creates a simulated information need for a user and functions as a platform against which situational relevance is assessed. In this study, the realism was taken into account by recruiting participants with genuine fiction reading interest and designing the simulated search scenarios based on previous knowledge in the common situations where fiction book search typically occurs. In the search scenarios, both researcher's control and proceeding according to participants' individual preferences were inherent. The experiment was conducted in a similar fashion with each participant, as the information given before and during the search scenarios were similar with each participant. As the experiment was similar with each test person, experimental control and realism was provided as recommended in the IIR evaluation model (Borlund 2003).

The participants conducted search scenarios in an experimental setting. Thus, it is not reasonable to generalize the findings to the population at large. Analyzing real log files could have produced more generalizable results. However, as fiction search in online environment has been scarcely studied previously, it was essential to produce basic information on readers' search behaviors and reveal possible differences between a traditional online library catalog and an enriched one. Thus, researcher's control was necessary to maintain in order to reliably compare readers' search behaviors between various search scenarios and catalogs. In addition, as the aim of this study was to understand fiction readers' search behaviors and search success in library catalogs, the measuring of search success had to be considered. For example, in analyzing real log files, the search success and users' criteria for selecting particular novels would have been challenging to detect. Observing, interviewing and analyzing users and their real search behaviors enabled the use of alternative performance measures such as search success and selected books' relevance as rated by users. As the search scenarios reflected the major browsing tactics for fiction retrieval, it could be assumed that the typical search approaches of fiction readers in natural setting would share some characteristics with our results.

When conducting the statistical data analysis in Studies II-V, the appropriateness of the applied measures was ensured. For the most part, the data was non-normally distributed which affected the selection of statistical data analysis methods. Non-parametric analysis methods were applied in situations where the non-normality prevented the use of parametric measures. As the variables measuring participants' search actions were manually encoded by the researcher, there were hardly any missing data. In addition, the outliers in the data were very few. In cases where outliers were detected, the statistical analysis was conducted two separate times: firstly, outliers included and secondly, outliers excluded. The conclusion was that the outliers did not affect the final findings.

The fiction book search as a research topic may be considered as an un-sensitive one. Thus, it is not likely that the participation in the study would have negatively affected the participants. In reality, the participants often commented on having enjoyed the experiment and gained new book discoveries during the search sessions. Before the experiment, the participants were informed that their on-screen activity and audio during the experiment would be recorded. Moreover, the participants were knowledgeable that their on-screen activities would be transformed in to quantitative variables, and the audio data anonymized in a sense that they could not be identified from the reported findings.

A few limitations should be noted. Firstly, the sample was biased towards females and the highly educated. Previous studies (i.e. Ross 2001; Tepper 2000) have shown gender to be associated to different reading interests of pleasure readers. For example, women are more likely to be heavy readers than males and they tend to read a greater variety of books than men (Ross 2001; Tepper 2000). In this study women's expected wider literary knowledge might have supported them to identify interesting titles better compared to men in the topical search. Then again, men's likely narrower reading preferences might have supported them to select interesting novels within a more limited range of authors and genres compared to women in the browsing task. By implication, males might have completed the browsing task quicker and with fewer search actions compared to women.

Secondly, as the questionnaire for the reading preferences in Study IV was used for the first time, the validity and the reliability of the instrument should be considered. Participants' answers to the questionnaire might have been influenced by factors such as time or mood. Also, participants' greater engagement in the book selection process in Sampo might have been a consequence of different reading preferences between the two groups. Even though the pre-questionnaire did not yield significant differences in reading activity or reading preferences between the two groups, it is possible that the users of Sampo were more interested in fiction reading than the users of Sata, which might have affected the results.

Thirdly, the limited number of participants might have influenced the final cluster solution in studying readers' search behaviors between reader groups in Study IV. It could be suspected that the cluster analysis would have yielded a solution with three to four clusters with a larger amount of participants. In addition, the limited number of participants prevented a simultaneous comparison of search behaviors between two catalogs and reader groups in Study IV. This could have had an association to further elaborate the results. Finally, as the systems used in the experiment were real, it was impossible to control participants' familiarity with particular books encountered during the experiment. This might have influenced the individual results. During the data collection, neither of the catalogs used in the experiment allowed users to add tags or book reviews to the book pages. The limited amount of user-generated content in the catalogs can be considered as a limitation in the identified interest criteria in Study V, as the influence on other users' tags and reviews of the book selection process cannot be evaluated.

4   OVERVIEW OF THE FINDINGS

This thesis consists of five studies, which approach fiction readers' search behaviors from different angles (see Figure 7). A genuine interest for fiction reading, lack of specific information need and a desire to find good books to read function as a shared background which ties together the different perspectives in the five studies. The studies have the following emphasis:

The Study I examines the typical search strategies for fiction in public libraries. The focus is on illustrating the association of demographic factors, reading activity and the type of books read with the use of major search strategies.

The Study II explores the basic characteristics of the fiction book search in various search scenarios in two different online public access library catalogs (OPACs). It reveals differences in readers' search behaviors in a traditional and an enriched online library catalog.

The Study III describes fiction readers' search actions and document viewing behavior during various query reformulation intervals (QRIs). The focus is on understanding how readers' search actions differ between successful and unsuccessful QRIs and which search actions predict greater search success.

The Study IV demonstrates the association of varied literary preference patterns and readers' search behavior in online library catalogs. Based on the findings, the study presents three search tactics for fiction books in library catalogs.

The Study V focuses on qualitative interpretations of the influential factors in fiction book selection in different search scenarios. It presents a five-dimension categorization of books' interest criteria and reveals a context-related pattern in readers' fiction book selections.
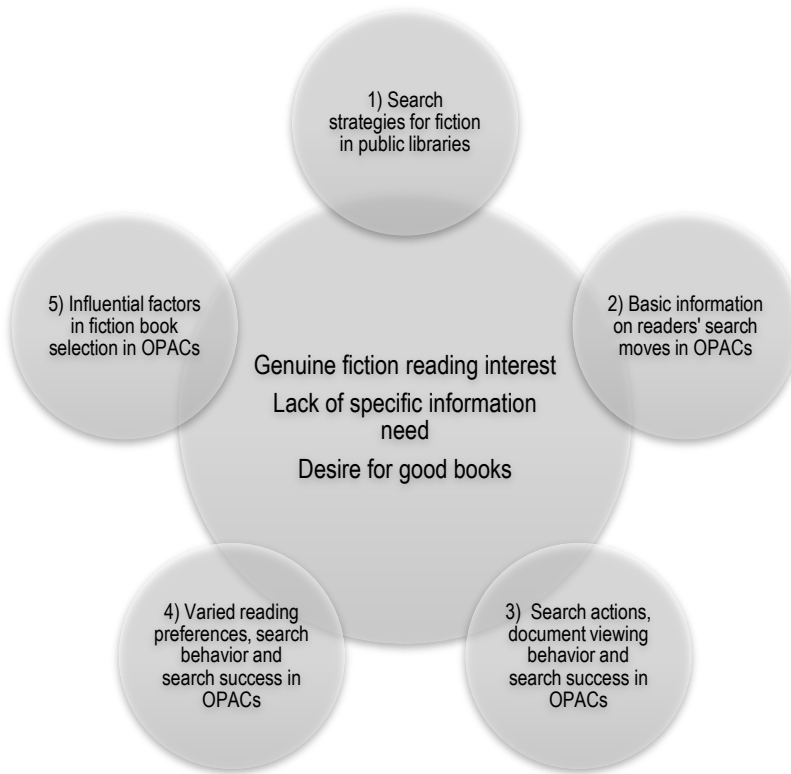
Figure 7. The themes of Studies I-V.

4.1   Search strategies for books in public libraries (Study I)

The Study I drew from the idea that various background variables would have an association with readers' book searching behavior while selecting books in public libraries. The central focus in Study I was in describing the typical search approaches for fiction books in public libraries and in illuminating the association of the demographic factors, the reading activity and the type of books read (fiction/non-fiction) with readers' search tactics for books in public libraries. The Study I made use of a survey collected in 2010 to study the outcomes of public libraries (see Vakkari & Serola 2012). The number of respondents to the questionnaire was 1000.

In the analysis, at first a one-way analysis of variance (ANOVA) was conducted to determine whether gender, an age, an educational level, a reading activity (the amount of fiction / non-fiction books read in the last 12 months) and the proportion of the fiction books of all books read in the last 12 months were associated with the use of different search tactics. For the second, in order to examine closely the association of the aforementioned variables, factor analysis and linear regression analysis was conducted. The aim of the factor analysis was to simplify the sequential analysis (linear regression analysis) in reducing the number of the dependent variables (search tactics variables).

The findings showed that the most popular search approach for books in public libraries was to search for books according to a known author or a title followed by browsing on the shelves and skimming the returned loans. Using a library catalog was the least popular search tactic. The ANOVA yielded some differences in the use of major search tactics between different demographic and reader groups, in particular in the use of three browsing tactics (browsing the shelves, browsing the book displays and skimming the returned loans). The tendency to browse was more common among females than males, and among active fiction readers compared to less active non-fiction readers. The known item search was used more by those with high education and those who read fiction compared to non-fiction readers. The tendency to search for books from a library catalog was more common among males than females. Asking help from a librarian was more common among respondents over 45 years old compared to younger respondents.

The multiple linear regression analysis yielded a clear association of gender, the educational level, the frequency of public library visits and the type of literature read with the use of search strategies. The variables affecting the use of browsing strategy were the frequency of public library visits together with gender, the amount of fiction read and age. The finding suggested that a public library visitor likely to select the books by browsing is an older female fiction reader, an active and experienced library patron having accustomed herself to the library's way of arranging and displaying books. Experienced library visitors are likely to having gained "behind the eyes knowledge" which helps them to use the browsing strategy. Experience is also likely to explain the fact that age was associated with the browsing so that older respondents were browsing more often than the younger. The Study I strongly supported previous findings (Ross 2001, Saarinen & Vakkari 2012) showing that significant association can be found between gender, the tendency to read fiction and the use of browsing strategy.

The variables affecting the use of the known item strategy were respondents' basic education and the amount of non-fiction read. The association of background variables with the use of the known item strategy can be elaborated twofold: firstly, respondent's basic education was a significant predictor of the known item strategy as highly educated applied it more often than respondents with lower education. Here, the association is likely due to the fact that people with higher education read more non-fiction than people with lower education (Griswold et al. 2005). Secondly, there had been some signs that in public libraries non-fiction is searched most by known item strategy (Timperley & Spiller 1999). The findings of this study are in line with previous research (Timperley & Spiller 1999). Moreover, the Study I indicated that the use of known item strategy is best explained jointly by basic education and the number of non-fiction books read. The more non-fiction read, the more often the known item search strategy was used.

The results of the Study I provided evidence that the use of different search tactics was associated above all with the frequency of public library visits, gender, an educational level and the type of books read. In the Study I, there was a strong association between the reading of fiction and the use of browsing strategy. Respectively, an association between the reading of non-fiction and the use of known item strategy was clear. Although the results confirmed and specified previous findings (Ross 2001, Saarinen & Vakkari 2012, Yu & O'Brien 1996), the variance explained by regression models remained low. Thus many of the factors associated with the variation of search strategies for books in public libraries remained unknown.

4.2   Search moves and success in library catalogs (Study II)

The Study II sought out to investigate fiction readers' typical search moves for fiction books and their association with search success in online library catalogs in different types of search scenarios. The center of attention in the Study II was in illustrating the possible differences occurring in readers' search moves and search success among different types of online library catalogs, a traditional one and an enriched one. A search move was defined as an identifiable basic action during the search

process (similar to work done by Bates 1990; Xie & Joo 2010). The Study II made use of participants' search logs recorded with Morae-software during the search sessions. The search moves applied in fiction searching in five search scenarios were analyzed. Firstly, the search logs were manually encoded and individual search moves were identified by examining the recorded user tests. Secondly, the identified search moves were measured by their frequency and duration (see Table 2). Success in search scenarios was measured by the ratings of the interesting novels found given by participants. The average book scores per task were calculated.

In general, participants were successful in finding interesting novels in both catalogs. In both catalogs, the highest book scores were given to the novels found in the three browsing tasks, where the participants proceeded according to their genuine interest. It seems that in both catalogs, the participants were the least satisfied with the novels found in tasks where author or topic was pre-decided by the researcher.

The findings showed significant differences in readers' search moves between library catalogs. The major differences were threefold:

user behavior due to variation in the access points to the collection in the catalogs,

user behavior on search result pages, and

user behavior on book pages.

In general, more effort (more search queries, search moves, and opened book pages) was needed in a traditional library catalog Sata to reach equivalent search success as in the enriched library catalog Sampo. The results demonstrated that in Sata, a typical search strategy for interesting titles seemed to be issuing queries and considering suitable entry terms carefully and devoting more attention to search results instead of book pages. In Sampo, time devoted to exploring the catalog's enriched front page and multiple entry points together with attention to the enriched result list seemed to be a common search approach. Also, searching without querying turned out to be a popular search approach, especially in browsing tasks in Sampo. When queries were issued in Sampo, they were formulated significantly more quickly than in Sata.

The capacity to retrieve an interesting result list was a crucial condition for successful discoveries in Sata, compared to browsing cover images, virtual book piles, and the enriched front page in Sampo. Hence, more emphasis was devoted to

querying in Sata compared to Sampo. While the single option for accessing the collection in Sata was querying, Sampo provided multiple entry points immediately at the front page. In Sampo, exploring the tag cloud, virtual book shelves, and the book cover carousel at the enriched front page offered an appealing starting point for the search, perhaps by causing curiosity, leading users to explore the catalog more closely. Moreover, the non-goal oriented scenario type and the lack of time pressure might have affected users' search behavior in Sampo in a sense that users were willing to explore all the system options before issuing any search queries.

Across all tasks, the users of Sata devoted more absolute time to search results compared to Sampo. In Sata, the search result list provided solely the author and title names whereas in Sampo, the result list was enriched, containing a cover image and a snippet of a blurb together with author and title names. The enriched result list in Sampo might have helped searchers identify possibly interesting titles more efficiently and more quickly than in Sata. With the rich information on the result list, the users of Sampo might have made pre-selections more quickly from the result pages, which may explain why the time invested in search results was shorter in Sampo compared to Sata.

Across all search scenarios, more time was devoted to book pages in Sampo compared to Sata. In Sata, information on novels' topics and themes was discovered on book pages by inspecting subject terms and publication information. A description and a cover image were provided with novels published recently, though in a separate window. Occasionally, the information on book pages was fairly limited, as subject terms were not provided. In Sampo, the book pages were always enhanced with many subject terms, a detailed description of the plot, a cover image, and automatic recommendations of similar titles. The book page display in Sampo may have intrigued the users more than in Sata, which may possibly explain the difference in time devoted to book pages between catalogs.

In Sata, the major difference between search scenarios occurred in time invested in querying. Queries were conducted significantly more quickly in the known-author task compared to other tasks. This suggests that participants were used to searching for novels with author names, while querying with themes or topics might have been more unfamiliar to users, requiring more time. In Sampo, queries were issued quickly across all tasks. As with Sata, the least time was invested in querying for the known-author task. Time devoted to the catalog's front page and time to pivot browsing seemed to be associated with the task type: users preferred pivot browsing and searched for clues of interesting titles at Sampo's front page more when the search

scenario was explorative by nature. When the author or the topic was pre-decided, the front page was given less attention.

The Study II explored the basic characteristics of fiction book searching in various search scenarios in online catalogs. The Study II revealed that fiction readers' search behaviors in a traditional catalog and an enriched catalog differ but may result in equal success. The results of the Study II offered a solid starting point for further examination of user behavior in online catalog fiction discovery.

4.3   Search actions and success during Query Reformulation Intervals (Study III)

The Study III arose from the premise of query reformulation intervals (QRIs), users' search actions that take place between two queries (Liu et al 2010). By examining fiction readers' search actions during QRIs, the study objective was to understand how fiction readers' search actions differ between successful and unsuccessful QRIs, which search actions are emphasized when interesting novels are selected, and which search actions predict the finding of very interesting novels compared to less interesting novels. The aim of the Study III was to examine fiction readers' search actions in *browsing situations*. Thus, the *"known author task"* was excluded from the analysis. Also, when conducting the user tests it was noticed that the tasks open ended browsing and searching without a query resembled each other. Thus, solely the tasks *"topical search", "open ended browsing"* and *"search by analogy"* were included in the analysis of this study.

Like in Study II, the Study III made use of the search log data. 80 user tests were manually encoded in order to identify different QRI types, and measure the search actions during the identified QRI types. In both catalogs three types of query reformulation intervals were identified: a) Successful QRI with very interesting novels selected (named as Successful QRI), b) Successful QRI with a little and somewhat interesting novels selected (named as Somewhat successful QRI) and c) Unsuccessful QRI with no interesting novels selected (named as Unsuccessful QRI). The search actions during QRIs were measured by their frequency and duration (see Table 2). Success in search scenarios was measured by the books' interest grading given by participants. Other factors determining success in a fiction book selection,

such as the aspects of a specific need, information / system literacy or collection coverage were also considered. However, at this point the data restricted the applying of these aspects in the measuring of search success in Study III.

In total, the data consisted of 1115 query reformulation intervals. The number of unsuccessful QRIs was significantly (p=.000) higher in the traditional library catalog Sata compared to the enriched library catalog Sampo. Significant differences were found in search actions between browsing tasks as searching for topical novels differed significantly from explorative browsing for interesting novels. The findings showed that when searching for topically relevant novels, most effort was required to assess somewhat interesting novels compared to very or non-interesting novels. In the topical search, the duration of QRI and dwell time on book pages were the longest during somewhat successful QRIs and the shortest during unsuccessful QRIs, i.e. the association was curvilinear. Deviating from the topical search, the results showed that in explorative browsing for novels, the duration of QRI and dwell time on book pages were linearly associated to document value. Selecting interesting novels required most effort, while non-interesting ones least effort in viewing results and book pages, but most effort in querying.

The results yielded that search actions during successful QRIs did not considerably differ between catalogs. Differences were found between catalogs in search actions during unsuccessful QRIs: in Sata query time and the number of viewed SERPs were significantly greater compared to Sampo. In Sata, an unsuccessful search pattern was detected: particularly at the beginning of each task searchers tended to end up going in circles, where much time was devoted to selecting suitable entry terms together with a quick glance over the results before issuing a subsequent query. When glancing at the results, they were noticed unsatisfactory, and book pages were not entered. In consequence, visits in SERPs were numerous together with long query time. In Sampo, a similar pattern was not discovered.

The Study III showed that the search actions that contributed to predicting the selection of more interesting novels a) varied slightly between catalogs but b) were similar between tasks within a catalog. Overall, the results demonstrated that in both tasks and catalogs the selection between interesting and not interesting novels was associated to few viewed SERPs and numerous opened book pages. Less viewed SERPs could be seen as a consequence of fewer queries. Thus, the first query should produce a satisfactory result list, from which to proceed to glancing at several book pages. It could be assumed that the more book pages are viewed, the higher the probability of encountering a very interesting one.

4.4 Reader characteristic, behavior and success (Study IV)

The Study IV drew from the idea that differences in fiction readers' literary preferences would have an association with readers' search behavior for fiction books in online library catalogs. The Study IV sought out to a) create a typology of fiction readers based on their responses to a questionnaire measuring literary preferences (Appendix II) and b) to understand how fiction readers with varied reading preferences are selecting interesting novels in online library catalogs. The precise focus of the Study IV was in revealing the associations between reader characteristics, search actions, search queries and search success in different search scenarios.

In Studies II and III, differences in participants' search behaviors were compared between the two catalogs. In the Study IV, the limitations inherent in the number of participants prevented a simultaneous comparison of 80 participants between two catalogs and two reader groups as it would have reduced the number of cases in each group overly low to conduct a reliable statistical analysis. Thus, the data from the two user groups was merged and the search behaviors were compared between reader groups. The Study IV made equal use of a questionnaire data, the manually saved search queries and the search log data. To identify groups of participants sharing similar responses on the reading preferences questionnaire, and to distinguish different preference patterns in fiction reading, a k-means cluster analysis was conducted. As a result of the cluster analysis, two different groups of fiction readers were elicited: *the aesthetics* and *the entertained*. The artistic and aesthetic pleasures of recreational reading were highlighted significantly more among *the aesthetics* compared to *the entertained*. *The entertained* particularly enjoyed escape and comfort that reading as a pleasurable activity offered.

The findings of the Study IV revealed significant differences in search behaviors between the two fiction reader groups. The major differences in search actions between *the entertained* and *the aesthetics* occurred in the number of search result page (SERP) visits and opened book pages, and in the time devoted to SERPs and opened titles. In "known author task", *the aesthetics'* familiarity with the given author's literary

production supported them to select more interesting novels with less effort compared to *the entertained*. Instead of opening several book pages, *the aesthetics* devoted much time to examining SERPs and tend to make the selections based on the information provided in the result list. *The entertained* clicked several items on SERPs to view the details, and made the selections based on the metadata provided on title pages. It seems that the result list provided enough information for *the aesthetics* to make the selections, whereas title page examination was more needed among *the entertained* to assess novels' interest level. This hints that a reader's literary knowledge influences search actions when selecting novels of a given author as knowledge of the author's literary production provides an advantage to identify more interesting novels compared to selecting titles of a previously unknown author.

In "open ended browsing", the reader groups were equally successful in selecting good books. *The entertained* often picked a familiar author's novels whereas *the aesthetics* preferred to search for previously unknown novels of a particular topic. *The entertained* identified interesting titles by few SERP visits and numerous opened titles whereas *the aesthetics* made several short SERP visits and opened few titles for closer examination. It seems that when browsing for good books, *the aesthetics* made the pre-selections from a larger group of search result snippets, but focused in detail on fewer book pages compared to *the entertained*.

In "search by analogy", search actions and search success were similar in both reader groups. The first query was often formulated to access the metadata of a previously read book for discovering suitable entry terms. Both reader groups made several SERP visits and opened many book pages in order to identify similarity between the titles. Both reader groups devoted more time to SERPs compared to book pages. Identifying familiar authors or titles, or appealing unknown titles were the key motives for entering onto book pages. In book pages, the similarities between titles were identified from metadata.

In total, the data for Study IV consisted of 770 queries, of which 445 were successful and 325 unsuccessful. In each task the queries contained typically two query terms. The query analysis revealed that the most popular search approaches in each search scenario were searching by an author or a topic in both reader groups. The findings showed that the majority of the search queries were one-dimensional containing solely a single aspect such as an author, a topic or a genre. When browsing for good books or novels similar to a previously read one, *the aesthetics* tend to find novels of a certain topic, whereas *the entertained* preferred unknown titles from familiar authors.

When initially coding the variables for search actions, particular patterns of search actions were detected. By watching on-screen activity, participants' actual search paths were monitored and each possible search action during the search process was detected in each scenario. Search actions were found to form typical sequences occurring in a similar fashion. The capacity of the identified search sequences to reflect natural user behavior could be argued. The search sequences could be considered reflecting users conditioning on the system and adapting to the interaction trails privileged by the system. However, they could also be seen as reflecting naturalistic search patterns. By integrating the information from the qualitative analysis of search queries and the participants' typical use patterns, it was possible to elicit three *suggestive* search approaches that give an insight into fiction readers' navigational behavior in library catalogs. Each of these search approaches were named as *a search tactic* which refers to a set of search moves that are temporally and semantically related (Wildemuth 2004). Three suggestive search tactics for fiction books were 1) Focused querying, 2) Topical browsing, and 3) Similarity based-tactic.

Figures 8 and 9 present two typical scenarios for the "focused querying -tactic".
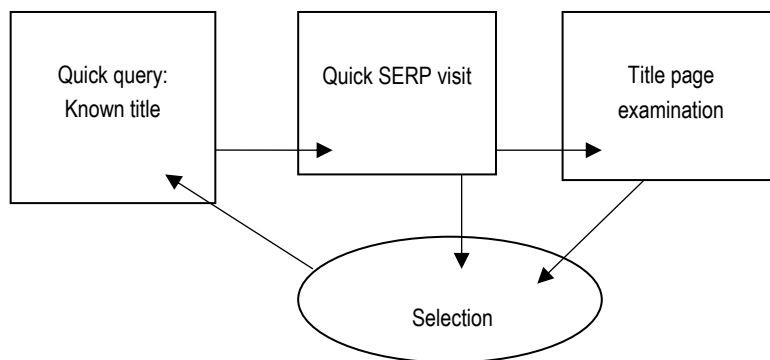
Figure 8.   Focused querying with known title.
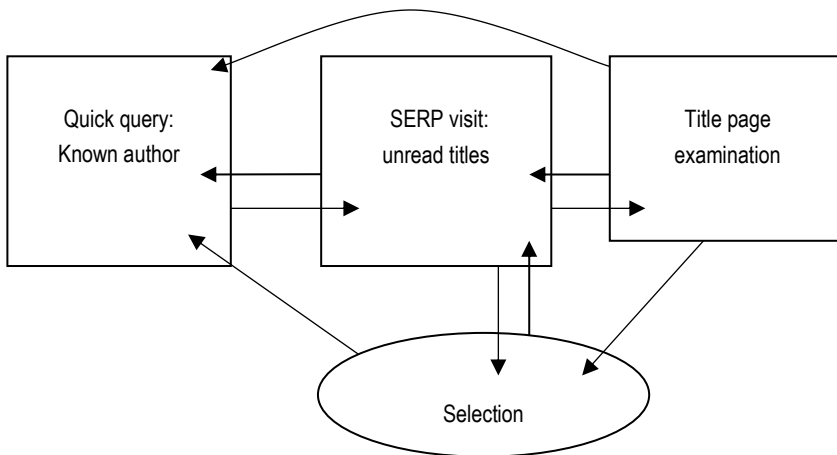
Figure 9. Focused querying with known author.

In "focused querying -tactic", searcher's literary competence and previous reading experiences function as a starting point for the search process. A searcher begins by issuing a quick query with either a previously known title or author. If a known title functions as an entry term, the first SERP visit is often quick (Figure 8). When examining the results, available metadata elements are the key factors for book selection: the richer the metadata is in SERPs, the more it supports quick book selections. If a known author functions as an entry term, the emphasis in SERP examination is on identifying unread titles (Figure 9). Again, rich metadata in SERPs supports book selections from the result list. If metadata is found to be limited, or if a searcher wants to examine an appealing title more in detail, the title page is opened. Title page examination results either in an engagement and selection, or in rejection followed by a new SERP visit or a subsequent query.

Typical search behaviors in "topical browsing -tactic" are provided in Figure 10.
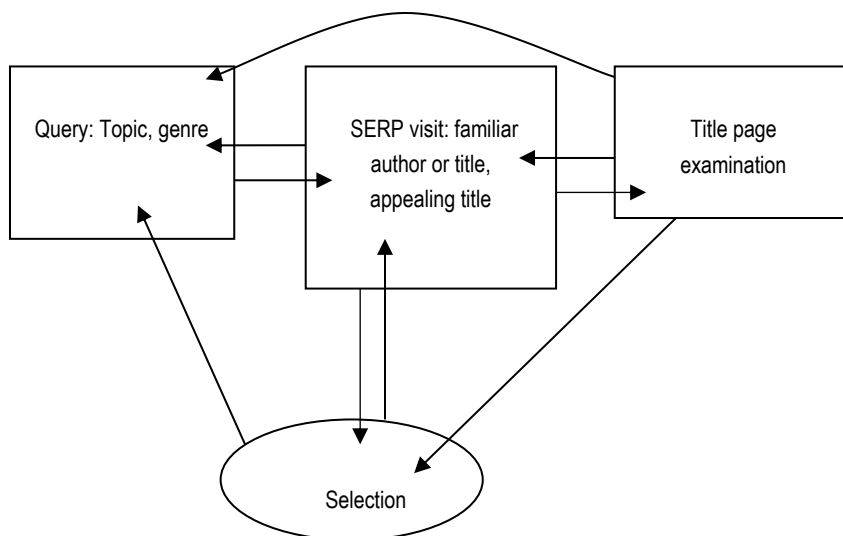
Figure 10. Search behaviors in topical browsing.

In "topical browsing -tactic", a starting point for the search process is a desired topic or genre of a novel. A searcher begins the search process by reflecting possible interesting topics or genres. In the beginning of "topical browsing", much time is devoted to formulating queries, as searchers aim at selecting unfamiliar and novel titles. SERPs are examined in detail in order to identify appealing titles or familiar authors. If the results are found unsatisfying, the searcher enters a subsequent query. If an interesting title is encountered, the searcher opens the title page for closer examination. In title pages the reading experience provided by the novel is assessed particularly by inspecting the subject headings and the blurb. If the content description provokes enough interest, the novel is selected. If the novel is rejected, the searcher returns to SERPs or issues a subsequent query.

Typical search behaviors in similarity based -tactic are provided in Figure 11.
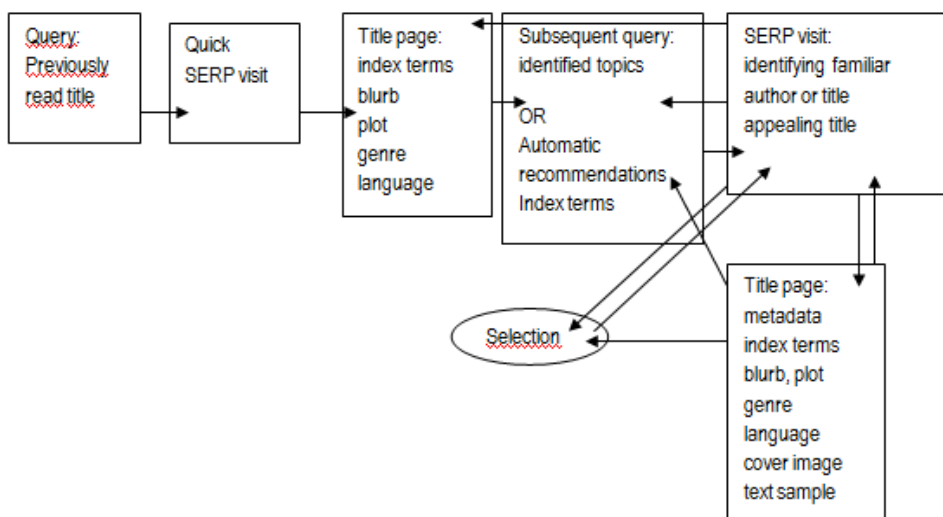
Figure 11. Similarity based tactic.

"Similarity based -tactic" is applied when a reader wishes to identify titles similar to a previously read interesting one. The starting point for the search is a previously read novel that has provided a good reading experience. The searcher begins by issuing a query with the previously read title as an entry term which is followed by a quick SERP visit. The title page is opened and the metadata is examined in order to identify the elements that are essential in the finding of new, similar titles. Usually, the searcher examines the index terms in order to identify the novel's topics and settings. The blurb and cover image are also paid attention to. Occasionally, the similarity is not associated to topics. Instead, an elaborate language or an author's style may function as the most desirable aspects in the novel. Thus, text samples (if provided) are also read. After identifying the desired features in the title page, they are expressed as a subsequent query. When examining SERPs, they are compared against the previous reading experience and similarities between novels are assessed. In title pages, the similarities are assessed particularly by inspecting the subject headings and the blurb. If the content description provokes enough interest, the novel is selected. If the novel is rejected, the searcher returns to SERPs. If SERPs are found unsatisfying, the searcher usually returns to the original title page for new

ideas for suitable entry terms. Alternatively, a searcher may rely on the features offered by the system to find similar items. Instead of issuing new queries, the search proceeds by following the system's features: in the title page index terms and automatic recommendations are examined and utilized to achieve new SERPs. If interesting titles appear in new SERPs or automatic recommendations, the similarities are again assessed by inspecting the subject headings and the blurb in new title pages.

4.5 Readers' interest criteria in fiction book search (Study V)

In the Study V, the focus for examining fiction book search behavior was shifted from users' search moves to their interpretations of the influential factors in fiction book selection. The prime motivation was to enhance the comprehensive understanding of fiction readers' book selection process in library catalogs by identifying the criteria for selecting a particular item in various search scenarios. The Study V sought to describe the differences in readers' interest criteria between different search scenarios and library catalogs, and in the applied interest criteria between rejected and selected novels. The Study V made use of participants' recorded interviews and conversations with the researcher during the search sessions. First, a five-dimensional categorization of fiction readers' interest criteria for examining the book selection process was presented; second, an in-depth study of the fiction readers' interpretations of the influential factors in fiction book selection in different search scenarios and library catalogs was provided.

The qualitative data contained a total of 931 selection mentions for novels, 607 from Sampo and 324 from Sata. Rejecting mentions were far less as the data included a total of 118 rejecting ones, 85 of them in Sampo and 33 of them in Sata. In analyzing the selection and rejection mentions for novels, five major interest criteria dimensions were elicited (Appendix III presents a detailed description of the dimensions): 1. Familiarity, 2. Bibliographic information, 3. Content, 4. Engagement and 5. Sociocultural.

The results show that when selecting novels in two library catalogs, the most applied interest criteria by readers were familiarity and bibliographic information.

Familiarity was related to known item search, as the majority of the selecting mentions in familiarity covered selecting novels from a known author or searching for a known title. When selecting a novel based on bibliographic information, the most important interest criterion was the title of a novel. In each task and in both catalogs, the title received a clear majority of the selecting mentions. In addition, the publication date and format of a book were also often mentioned. The format of a book was associated to a short story as a form of literature, as the mention of "short story" in bibliographic information either caused an immediate rejection or a book selection with delight. The difference in participants' reactions towards short stories was mostly due participants' personal literary preferences concerning various literary types, like opinions for poetry or plays as a recreational reading material.

The content dimension included selecting mentions related to the intellectual content description provided by library professionals such as subject headings, description of a novel's plot (blurb) and genre classification. The engagement in reading as an interest criterion refers to the mentions indicating general liking or disliking of a novel. Readers often reflected the ideas and expectations triggered by a particular novel and made the selections based on these preliminary estimations. Mentions such as "This sounds really interesting" and "I would choose this book for real" were common in the engagement dimension. Sociocultural aspects were the least mentioned in readers' book selections compared with the other major dimensions. Most of the selecting mentions in sociocultural dimension were related to identifying a personal connection to a particular novel which triggered an interest to select the book. The personal connection might have been identified from a title of the novel or by examining the blurb and the subject headings on a title page.

In both catalogs, the rejecting mentions were most often associated to bibliographic information such as an author in a sense that readers had a clear vision of the undesired authors. Also, publication date was mentioned as readers often avoided older books. The rejecting mentions were also associated to a single subject term on a title page: an undesired theme or topic was easily detected from the subject headings section and the novel was quickly rejected.

The findings show that when the author was given, readers selected the novels mostly based on bibliographic information, content description and with previous knowledge on the given author's literary production. The major difference compared with the other search scenarios was the notable influence of a novel's title in the book selection in the known author search. When the search results were examined, the titles were not just passively skimmed through. Instead, the participants analyzed and interpreted the ideas provoked by the titles in detail and evaluated the possible

reading experience offered by a particular title. An appealing title was often mentioned as containing an aspect close to one's personal life. Also, the title could have been mentioned as engaging if it was funny, included interesting concepts or a clever play on words. Overall the results show that when selecting books from a known author, bibliographic information turned out to contribute more to the book selection compared with the content description. The content description was often used as a secondary criterion for the selection after detecting an appealing title from the search results.

In topical search, familiarity with the given topic was a major factor in the book selection. Participants used their knowledge of authors and titles in two ways when selecting topical novels: they issued a keyword search with a suitable author or title; or they issued a keyword search with a topic (e.g. upper class nineteenth century) and selected known items from the result list. Compared with the other search scenarios, topical search was the only one where content description overcame bibliographic information in selection. Both in Sampo and Sata, readers detected the topical relevance of a novel from the subject headings and the blurb on an item page. The results show that the enriched content description in the search results and item pages was the dominant interest criterion over the familiarity in Sampo. In Sata, the participants relied more on their knowledge of literature over examining the content description which was often limited and narrower than the one in Sampo. The results also show that bibliographic information had a great role in selecting topical novels in both catalogs. The overwhelming majority of those mentions were about the title of a novel. Differing from the other search scenarios, in topical search title mentions did not refer to particularly appealing or likable titles. Instead, they focused on identifying classics such as Jane Eyre or Pride and Prejudice representing upper class life in the nineteenth century.

In the open ended browsing task, there were no differences in selection criteria between catalogs. Readers in both catalogs relied greatly on their literary knowledge and familiarity with preferred authors and titles. A common pattern to select "just good books" was to search for favorite authors' novelties, or titles from personal reading lists kept and updated constantly on mobile devices and personal organizers. Selecting novels based on familiarity was detected at the beginning of the search scenario: if a participant was unable to come up with an author or a title after reading the indicative request, the task was experienced as challenging and difficult. Literary awards, television shows, newspaper articles and radio programs concerning recently published fiction books were mentioned as influencing the book selections in the open ended browsing task as they provoked ideas for the keyword searches. Since

participants' literary knowledge functioned as the most frequent interest criterion in the book selections for "just good reading," less emphasis was given to bibliographic information and the content description. When mentions concerned bibliographic information, the title of a novel was again the most important criterion. Engagement in reading, especially a positive reading experience provoked by a novel, was also perceived as an important selection criterion in the open ended browsing task. This differs from the known author search and the topical search scenarios and hints that when readers are selecting novels based on their true reading interests, the reading experience provided by a particular title is valued more compared to selecting topical novels or novels of a given author.

In the search by analogy task, the participants selected three similarly interesting novels to a previously read one. At the beginning of the search scenario, the participants interpreted the idea of similarity and applied this interpretation to their book selection process. Participants mentioned the similarity to be related to the following aspects in novels: the same genre, the same author, a similar topic, the same mood and a similar narrative style. Depending on the aspect of similarity, the search scenario was perceived either as an easy or a challenging one. If similar novels were selected based on a genre or an author, the search scenario was often completed easily. If a similar topic, mood or a narrative style was the starting point for the selection, the participants had difficulties in discovering similarly interesting novels. The similarity was often mentioned being a combination of various aspects, such as a novel with a particular narrative style covering similar topics. These expectations were rarely met and participants compromised between similar topics and interestedness. In the search by analogy task, the most important criterion for the book selections in both catalogs was bibliographic information, particularly the title and the author of a novel. Selecting similar novels based on an author or a title often resembled a topical keyword search: participants issued a query with the wished for topic of a novel, and selected novels from the search result list based on the author information and the appealing title. The content description and familiarity with authors and titles in selecting similar novels was emphasized slightly more in Sampo than in Sata. Similarly, the users of Sampo mentioned engagement in reading notably more compared to the users of Sata. It seems that in a traditional catalog, selecting similar novels was associated to detecting an appealing title after topical search, whereas in the enriched catalog the title, the content and the positive expectations provoked by the former were emphasized evenly.

The findings from the four search scenarios hint that in library catalogs, the process resulting in a book selection differs from the one resulting in a book

rejection. Overall, it seems that the rejection of a book is a simple and quick process in which a single metadata element (such as a subject term, a publication year) often determines the decision of a non-interesting book and a rejection. Quite the contrary, a book selection seems to be a complex chain of search actions, during which the influence of previous knowledge, expectations, emotions and interpretations is associated to the outcome.

5 DISCUSSION

The primary motivation for this thesis was to contribute to the comprehensive understanding about search behaviors for casual-leisure reading material and fiction readers' book selection process in online library catalogs. The five original research papers and the summary illuminated various factors associated with search success when searching for novels in digital environments, which is essential as digital fiction collections continue to grow. They enlightened readers' typical search approaches for books in various search scenarios and described the differences in user behavior between different library catalogs. The first part of this chapter summarizes the results and contributions of the five subsidiary studies (Studies I–V). After this, the findings are discussed in depth threefold:

For the first, the fiction book search process is discussed by combining the major findings of five subsidiary studies (Studies I-V). At the beginning of this section, a model describing the fiction book search in digital environment is presented. The focus is on revealing how fiction readers' book search and selection appear in the light of information gained in this thesis. As fiction book search in digital environments was scarcely studied previously, there were no established theory or practice to adapt in the design of the thesis. Instead, the thesis can be considered as a basic research suggesting an approach for studying search behaviors for casual-leisure readings.

For the second, the findings of this thesis are discussed from a wider perspective of casual-leisure search, enlightening the differences and similarities between a search for recreational reading material and a search for non-fiction information resources. Moreover, the contribution to the research methods applied for examining casual leisure information behavior is discussed.

For the third, the findings of this thesis are elaborated from a viewpoint of a system design. Implications for practice in public libraries' catalog development are presented.

5.1 General summary of findings

This thesis is positioned at the intersection of information science and sociology of reading. It contributed to our understanding about fiction book search in information systems as a context-related process. It explored genuine fiction readers' approaches for selecting recreational reading materials mediated by two different online library catalogs. Deviating from the main-stream of studies in information retrieval, it concentrated on the search scenarios which lack a specific information need or a task to be solved. Instead, the thesis drew from an idea that fiction book search occurs with a call for "just good books". The thesis made use of the IIR evaluation model designing simulated search scenarios (Borlund 2003) for casual-leisure readings and applying alternative performance measures in the evaluation of fiction book search. This approach to studying the selection of recreational readings in a large scale controlled user experiment is innovative. The thesis revealed the typical search approaches for fiction by eliciting the navigational patterns of fiction readers while selecting novels in library catalogs. The thesis enlightened the associations of search behaviors and search success in different search scenarios. The thesis illuminated fiction readers' interest criteria for novels, typical preference patterns in fiction reading and described the association of various preference patterns with the book selection process in online library catalogs.

Study I served as a general introduction to the methods readers' apply in their book selection. The study drew from the idea that various background variables would have an association with readers' book selection behavior while selecting books in public libraries. The results provided evidence that the use of different search tactics was associated above all with the frequency of public library visits, gender, an educational level and the type of books read. In this study, there was a strong association between the reading of fiction and the use of browsing strategy. Respectively, an association between the reading of non-fiction and the use of known item strategy was clear.

Study II focused on the basic characteristics of fiction book searching. It looked in fiction readers' typical search moves for fiction books and their association with search success in online library catalogs in different types of search scenarios. The study revealed that fiction readers' search behaviors in a traditional catalog and an enriched catalog differed but resulted in equal success. The differences were bound

up with user behavior due to variation in the access points to the collection in the catalogs, user behavior on search result pages and user behavior on book pages.

Study III continued elaborating the differences in user behavior while selecting novels in online library catalogs. It looked in more detail user behavior during query reformulation intervals with an attempt to understand which search actions predict the finding of very interesting novels compared to less interesting novels. The study showed that the selection between interesting and not interesting novels was associated to few viewed SERPs and numerous opened book pages. Also, the study showed that the task type was associated to readers' document viewing behavior.

Study IV explored differences in fiction readers' literary preferences and their association with readers' search behavior for fiction in online library catalogs. The study revealed a two-dimensional typology of fiction readers based on their responses to a questionnaire measuring literary preferences. It showed significant differences in search behaviors between the two fiction reader groups. The major differences in search actions between *the entertained* and *the aesthetics* occurred in the number of search result page (SERP) visits and opened book pages, and in the time devoted to SERPs and opened titles. Based on the findings, the study presented three typical search tactics for fiction books in digital environments: 1) Focused querying, 2) Topical browsing, and 3) Similarity based-tactic.

Study V examined the influential factors in fiction readers' book selection. It created a five-dimensional categorization of readers' interest criteria for novels and evaluated the criteria for selecting a particular item between search scenarios and library catalogs. The findings showed a book selection process as a complex chain of search actions, during which the influence of previous knowledge, expectations, emotions and interpretations was associated to the outcome.

5.2   Fiction search as a context-related process

In general, the results of this thesis both confirmed and extended previous knowledge on fiction book selection. As noted already in Spiller (1980) and Yu & O'Brien (1996), *the known item search* and *browsing* continued to bear significance as the most popular search strategies for fiction books both at the shelves of public libraries

and in digital environments (Pejtersen 1989). However, this thesis extended previous knowledge on fiction book search by revealing a number of factors associated with the use of the major search strategies and illuminating the specific navigational patterns for the known item search and browsing in digital environments. Overall, the findings showed a context-related pattern in readers' fiction book selections in a digital environment. A combination of readers' situational reading interest, search capacities, "behind the eyes" knowledge, affective factors and a well-functioning interaction with a system used was shown to result in a successful book selection. Based on the findings, a general model for a fiction book selection in a digital environment is presented in Figure 12.
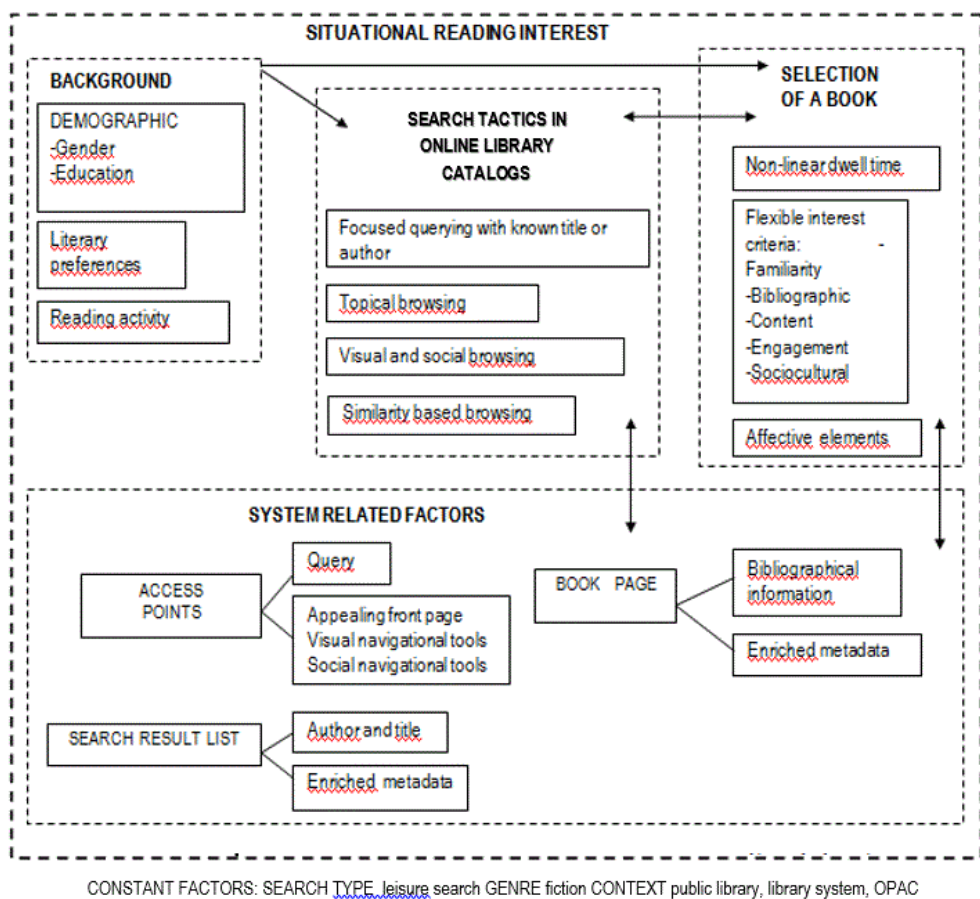
Figure 12. The model of a fiction book selection in a digital environment.

The model (Figure 12) suggests a fiction book search to consist of five interconnected parts which each affect the overall search process: 1) the situational reading interest, 2) the background variables, 3) the typical search tactics in online library catalogs, 4) system related factors and 5) the factors that influence the selection / the rejection of a particular item. In the model, the factors that are held constant in each individual fiction book search are the search type as a leisure search, the fiction genre and the search context as a public library and an OPAC connected to the library system.

The situational reading interest. Above all, the results of the thesis appear to suggest that the search for fiction occurs in a particular scenario, whether it to be discovering novels for a holiday trip or searching simply something good to read. The search scenario creates the situational reading interest and functions as the reference point against which the continuous stream of the search results and hints of possibly interesting readings are evaluated during the entire search process. The findings of the thesis showed that according to the situational reading interest, different expectations towards the novels to be selected were created and the novels found and encountered were assessed in the light of these expectations. Hence, it appears that readers' selection criteria for novels and their typical navigational and querying patterns varied by the situational reading interest. The interest criteria for novels created in Study V was applied in a flexible and multiphase way depending on the search scenario and phase. As a result, the book selection process turned out to be a combination of affective, personal, socio-cultural and metadata-related factors.

Overall, the results of this thesis suggest a tendency for the book selection process to be initiated by figuring out ideas for keyword searches such as known authors, titles or books nominated for a literary award across all search scenarios. This confirms that searching for books based on authors' names (Spiller, 1980; Yu and O'Brien, 1996) applies also in online library catalogs. At the beginning of a search scenario, a failure in discovering ideas for interesting reading based on literary knowledge appears to be a major obstacle in the book selection process.

It would appear that the search scenario affects the search behaviors for fiction books in a variety of ways. If searching for books from an unfamiliar author, the first phase of the selection seems to be browsing for an appealing title. After that, the second phase is to search for novels published recently followed by examining the content description to identify an interesting item. If the author's literary production is previously known, the first stage appears to be selecting the most recently published title or well-known favorites to be re-read. After that, the next stage is to

select engaging titles followed by discovering interesting themes from the content description.

Searching for novels according to a topic was experienced as a not-likely means of engaging oneself in to the fiction book selection process by the participants. However, in selecting topical novels, an important starting point for the search was participants' prior literary knowledge which was used in formulating suitable queries and in identifying suitable titles from the search results. The results of Studies III and V show that topical search differs from the other search context so that together with familiarity, a dominant interest criterion for selecting a particular item is the content description instead of bibliographic information and engagement in the reading. At Study V, during the selection of topical novels, participants, primary selection criterion was novel's topicality, which was detected from professional metadata in the content description. After selecting three topical novels, participants' assessed their interest level. As the topic was sometimes perceived as non-interesting by the participants, also the novels selected received sometimes ratings as only a little interesting. Thus, interest towards a novel was a secondary selection criterion for topical novels. Emphasis on the content of the novels and the irrelevance of liking in participants' interest criteria suggest that the selection of topical novels resembles that of non-fiction documents. Like in Gwizdka (2014), the results of Study III yield that when searching for topically relevant novels, most effort is required to assess somewhat interesting novels compared to very or non-interesting novels. Similarly to Vakkari et al. (2014) it seems that when browsing for topical novels, in *"selection stage most time is used for judging borderline cases compared to more valuable or valueless cases"*.

Differing from selecting known items and topical novels, selecting interesting novels in explorative browsing task appears to be a complex and a multi-phased process where a) personal, affective and situational factors bear the most significance to the book selection; b) serendipity is continuously involved as the readers rarely know what they wish to discover; and c) the satisfying result is challenging, even impossible to predict at the beginning of the search scenario. Similar phenomenon has been noted by Koolen et al. (2015) who found that book requests for social book search were often a combination of content, context and examples. The results of Study III suggest that deviating from topical fiction search, in explorative browsing selecting interesting novels requires most effort and in the selection stage, most time is used for examining the most interesting opened book pages. The finding is in line with Liu & Belkin (2010) and Liu et al. (2012) who found that when selecting non-fiction documents, dwell time on document pages was linearly associated to document usefulness. The observation of participants revealed that in the topical

search, there appears to be a tendency to select first a novel with high topical value and then assess its interest level. Thus, interest level is often a secondary criterion for selection in the topical search, whereas in the browsing task searcher's genuine interest is a strong motive to select a certain novel. It seems that identifying truly interesting novels in open ended browsing situations requires long dwell time on book pages, while eliminating not interesting ones requires least effort.

The results of Study V suggest that selecting fiction books with similar reading experience in library catalogs relies greatly on novels covering similar topics. Readers in Study V often settled for a topical similarity yet a mood or a particular reading experience might have been the primary criterion for the similitude. Sometimes, recommender systems fail to create recommendations based on the affective side of books. Thus, it would appear that novels providing similar reading experiences to a previously read one are challenging to detect in library catalogs.

Overall, the results from Studies II-V illustrate the explorative browsing for novels in a digital environment as a challenging search context as it does not provide a particular, clearly defined starting point for the search. Nevertheless the difficulty often experienced at the beginning of the search scenario, readers seem to aim at selecting truly engaging titles by applying their literary capacity in the book selection process.

The background. Within the search scenario, the motivation for the fiction search arises from a reader's background. The results from Study I show that demographic factors such as gender and educational level affect the selection of a search strategy for books in public libraries. It appears that gender is associated with the use of browsing tactics as female participants in Study I reported more browsing behavior compared to males. This association is most likely due to the females' greater tendency to read and especially to read fiction, as shown previously (Griswold et al. 2005; Moyer 2007; Ross 2001; Tepper 2000). There had been signs that the use of browsing strategy would be associated with the amount of fiction books read (Ross 2001; Saarinen & Vakkari 2013). For example, Saarinen & Vakkari (2013) showed that avid fiction readers used browsing search strategy more often than occasional readers. The Study I strongly supports these findings suggesting that significant association can be found between gender, the tendency to read fiction and the use of browsing strategy. Similarly, educational level has been found to be a major factor affecting leisure time reading as people with high education read more than people with lower education (Griswold et al. 2005; Kraayakamp & Dijkstra 1999; Ross et al. 2006). The results of the Study I show that the use of known item strategy appears to be associated with gender, an educational level and the amount

of non-fiction read. Previously, there had been some signs that in public libraries non-fiction is searched most by known item strategy (Timperley & Spiller 1999). The findings of the Study I are in line with previous research and indicate that the use of known item strategy is best explained jointly by gender, basic education and the number of non-fiction books read. Therefore, it would appear that even before the book search process begins, the varied reading preferences among different demographic groups generate preconditions and different expectations towards the books to be selected.

However, the demographic factors do not alone determine everything. Even greater impact to the fiction book selection appears to fall to person's motives for fiction reading, reading activity and to one's literary preferences as in genres, literary styles and favorable authors. As in previous studies (i.e. Miesen 2003; Ooi & Liew 2011; Saarinen & Vakkari 2013), the participants in the Study IV enjoyed the entertainment, escape, aesthetic pleasure and utilitarian aspects that reading as a pleasurable activity offered. The Study IV elicited different preference patterns in fiction reading and created two fiction reader types: *the entertained* and *the aesthetics*. This typology for fiction readers is in line with Saarinen & Vakkari (2013), where fiction readers were divided into entertainers, esthetes and realists. The Study IV revealed significant differences in search behaviors according to one's literary preferences.

The findings of the Study IV suggest that when readers are selecting novels from a known author, the search behaviors vary according to one's familiarity with the author's literary production. A high familiarity with the particular author's literary production appears to support high search success with less effort in comparison with a limited knowledge on the author's literary works. In particular, a high familiarity seems to support a reader to make a book selection based on the information on the result list, whereas a more limited knowledge requires a reader to open several title pages in order to make the selection based on the metadata provided on the title pages. This hints that a reader's literary knowledge influences search actions when selecting novels of a given author as knowledge of the author's literary production provides an advantage to identify more interesting novels compared to selecting titles of a previously unknown author.

The results of the Study IV show that in browsing scenarios, the search behaviors vary according to one's literary preferences but result in equal search success. It would appear, therefore that in browsing for good books to read, readers apply their personal interest and previous literary knowledge to their search behaviors in a successful manner. Differences in search behaviors according to one's literary

preferences are related to the search queries and to the user behavior on search results and book pages. The findings of the Study IV showed that when browsing for good books or novels similar to a previously read one, *the aesthetics* tend to find previously unknown novels of a certain theme, whereas *the entertained* preferred unknown titles from familiar authors. This is in line with Saarinen & Vakkari (2013) who noticed that reader's literary competence was associated to browsing behavior when selecting novels: esthetes with interest in high standard novels and a broad variety of novels accessed books by open-ended browsing more compared to participants reading for entertainment. Participants reading for entertainment knew in advance the titles they were seeking and open-ended browsing was not a common search strategy for them. Like in Tang et al. (2014), it would appear that *the aesthetics* in the Study IV were willing to explore novel book recommendations and expand their reading horizon, whereas *the entertained* with high knowledge of their reading preferences preferred an author browsing -strategy.

The findings suggest that different preference patterns for fiction books are clearly associated to readers' search behavior, i.e. the number of viewed SERPs, opened book pages, dwell time on book pages and the type of search queries. In general, *the entertained* identified interesting titles by few visits to search result pages and numerous opened titles whereas *the aesthetics* made several short search result visits and opened few titles for closer examination. It seems that *the aesthetics* made the pre-selections from a larger group of search result snippets, but focused in detail on fewer book pages compared to *the entertained*. Overall, the Study IV implies that searchers' literary knowledge and previous reading experiences are associated to the selection of a suitable search approach and an interesting book in each search scenario: at all times the starting point for the search appears to be either a known title or an author, an interesting topic or genre or a previously read well-known title. The new works of literature are assessed in relation to the earlier reading experiences.

Search tactics. The Study IV presented three suggestive search tactics for fiction books in digital environments: 1) Focused querying, 2) Topical browsing, and 3) Similarity based-tactic. By integrating the information from the five subsidiary studies, a visual and social browsing is added in to the common search approaches for fiction in digital environments (Figure 12). The findings of the Study IV indicate that readers' typical search behaviors and the factors influencing the book selection are different in the identified search tactics. The typology of the search approaches in the Study IV shares similarities to the work by Pejtersen (1989). However, the search tactics identified in the Study IV give a more detailed insight in to reader's

navigational patterns presenting the actual search moves and the content of the search queries during each of the identified search tactics.

The selection of a suitable search approach appears to be a combination of personal and system related factors: as aforementioned, the background and the previous literary knowledge affect the expectations towards the novels to be selected, and by consequent to the well-functioning search approach. The system to be used affects the search approach selection as it creates the information environment where the search process occurs either restricting or enabling versatile user behavior. User's previous experiences and skills in navigating through an online environment in a quest for good books have a strong association with the flexibility in selecting suitable search approach according to a search context.

The findings revealed a detailed navigational patterns for fiction books during the identified search tactics. Like in Oksanen & Vakkari (2012), in general the fiction book search process is a combination of search queries, search results inspection and a book page examination. The identified sequences of search moves and the analysis of the search queries within the identified search tactics showed that readers' alter the use of a search approach in a flexible manner according to the search context. In all, it appears that the "focused querying" is an effective and a well-functioning search tactic for novels regardless of the search scenario. However, previous knowledge of authors and titles is required in order to apply the "focused querying -tactic". An enriched catalog might support "focused querying" more over a traditional one, as ideas for authors and titles are often provided already at the entry page of an enriched catalog. Then again, the "thematic browsing -tactic" seems to be a well-functioning search approach for novels particularly in scenarios such as topical search and open ended browsing. Ability to reflect and identify desired themes and issue them as query terms is required to apply the thematic browsing tactic. In addition, ability to identify appealing titles from search results presenting topical novels is essential in identifying interesting titles effectively.

The two similarity tactics seem to be capable particularly in tasks where the searchers proceed according to their genuine reading interests. These tactics differ from "thematic browsing" especially in a sense that the departure of the search is more limited: a searcher has in mind an appealing title, which functions as a reference point of SERPs. The query terms are identified based on the previously read or already found interesting title, and the results are examined in relation to the known or already found title. The visual and social browsing seem to be suitable approaches for novels especially in an explorative search where a reader has only an ill-defined wish for good readings. The social browsing according to other users

recommendations and visual browsing according to the system's features (i.e. book cover browsing) requires an ability to identify an appealing title from a continuous stream of possible interesting readings. At its' best, the social and visual browsing provide great possibilities to serendipitous book discoveries, enabling searching without a query.

Extending previous knowledge concerning fiction book selection, this thesis describes the process for selecting a fiction book as a dynamic one during which the identified search approaches are altered flexibly. The query analysis in the Study IV revealed that the most popular search approach in each search scenario was searching by an author, especially at the beginning of a search. A tendency to begin the search process with a query by a known author or a title might be considered as a consequence of the experimental design. The participants may experience it as a necessity to succeed in their searchers, which is why they might prefer a straightforward and simple approach with favorite authors and titles. Nevertheless, in studying complex search queries for books in social media a similar pattern has been detected: Koolen et al. (2015) noticed that two dominating aspects in complex search queries for books in social media were the content of the book and the familiarity of the desired reading experience. Familiarity was described as searching for similar books as known ones or books related to a previous reading experience. It would appear, therefore that a common pattern for selecting fiction books is to begin the search with familiar authors and titles. However, there also appears to be a tendency to alter the search tactic during the search process within a search scenario. To exemplify, the first novel might be discovered with a focused querying. After that, the search approach might be altered to a visual browsing and then again to the similarity based tactic. That is to say, ideas for the consequent searchers are developed during the search process and searchers alter the suitable search approach according to the new ideas. This resembles Bates' idea of berrypicking, where each new piece of information that one encounters provides potentially new ideas and direction to follow (Bates 1989). Furthermore, when users familiarize themselves with the system to be used, they seem to become more courageous to endeavor different features and more willing to explore alternative approaches for novels.

System related factors. In the model of a fiction book selection in a digital environment (Figure 12), the system related factors are described as having an effect to the search tactic selection and to the selection of a particular item. In a digital environment, the system creates the actual information environment in which the search process occurs. In the fiction book search, both the search tactic selection and the book selection are bound to the system's capacity to provide various access

points to the collection and to the metadata elements provided in the search results and on the title pages.

In a traditional catalog that does not support browsing with visual or social navigational tools, a query is a necessity for browsing the collection. An essential precondition for a satisfying query is a knowledgeable literate user who is capable of inventing suitable author or title names or desired themes as keywords. In Studies II and III, the capacity to retrieve an interesting result list was a crucial condition for successful discoveries in the traditional catalog Sata, compared to browsing cover images, virtual book piles, and the enriched front page in the enriched catalog Sampo. Hence, more emphasis was devoted to querying in Sata compared to Sampo.

While the only option for accessing the collection in Sata was querying, Sampo provided various visual and social browsing features at the start page of the catalog. This might explain why time devoted to the catalog's front page seemed to be a typical search approach in Sampo. This is in line with the finding by Anfinnsen et al. (2011) that tag clouds encouraged users to browse a catalog in depth and enhanced active user involvement. Buchanan and McKay (2011) noticed that when readers were expressing their reading wishes in query terms, lack of knowledge was sometimes an obstacle to successful discoveries. Thus, the social and visual navigational tools that enable browsing without querying might have increased the possibility of serendipitous discoveries in Sampo. As Hassan-Montero and Herrero-Solana (2006) point out, querying requires the user to formulate information needs, whereas visual browsing allows the user to recognize information needs while scanning the interface.

In Sata, the identification of an interesting title from a search result page might have been challenging for users, as only author and title names were provided. This was even harder when the author was previously unknown. In Sampo, the result list was enriched, containing a cover image and a snippet of a blurb together with author and title names. Across all tasks, the users of Sata devoted more time to search results compared to Sampo. The enriched result list in Sampo might have helped searchers identify possibly interesting titles more efficiently and more quickly than in Sata. This agrees with Vakkari and Pöntinen (2015), who found that an enriched catalog helped users identify potentially clickable items on the result list sooner as compared to a traditional catalog. Also, enriched result list might draw searchers into book pages for more detailed examination, and support the identification of interesting novels. Moreover, in book pages automatic recommendations are generated and searchers may proceed without entering a subsequent query. In the traditional catalog the result list provides titles and author names. If they are found not interesting, a necessity for

proceeding is a re-formulated query. This may explain the difference in search actions during unsuccessful query reformulation intervals between the catalogs.

Similar to work by Pöntinen and Vakkari (2013), Studies II and III found that more time was devoted to book pages across all tasks in Sampo compared to Sata. In Sata, information on novels' topics and themes was discovered on book pages by inspecting subject terms and publication information. A description and a cover image were provided with novels published recently, though in a separate window. Occasionally, the information on book pages was fairly limited, as subject terms were not provided. In Sampo, the book pages were always enhanced with many subject terms, a detailed description of the plot, a cover image, and automatic recommendations of similar titles. The book page display in Sampo may have intrigued the users more than in Sata, which may possibly explain the difference in time devoted to book pages between catalogs. Users of Sata opened more book pages to select equally interesting titles as compared to users of Sampo. This might partly be a result of Sampo's enriched result list, as opening book pages was not necessary to identify appealing titles.

The results of the Study V suggest that interest criteria in selecting novels varied by catalog. These differences concerned familiarity, bibliographic information, content description and engagement in reading. In general, the readers in Sata emphasized the association of prior knowledge and bibliographic information in selecting novels, whereas the readers in Sampo highlighted the influence of content description and engagement in reading in their book selections. As a traditional online library catalog, Sata offered little support for readers to select novels other than known items. The users of Sata were to rely on their own ideas and literary competence for good reading as the catalog did not provide any book recommendations, lists of novel publications or diverse navigation features. As commented on by the participants, topical keyword searchers were perceived as an unnatural search tactic and metadata was often limited and displayed in an unappealing way. Lists of subject terms were often the single attribute providing information for creating the image of the expected reading experience. This explains why familiarity and bibliographic information (the title) were the most influential factors in readers' book selections in Sata. In Sampo, the enriched features at the starting page, in search results and at item pages distinguished the book selection process greatly from Sata. Sampo offered a continuous stream of possibly engaging reading via visual and social navigational tools, which may have altered the search process as more serendipitous compared to relying on known authors and titles. This may reduce the role of the familiarity in book selections. The enriched item pages

with cover images, blurbs and automatic recommendations might have provided more complementary information to readers for generating more associations of the novels compared with Sata. As a consequence, a general interest toward the novels selected might have been easier to detect in Sampo than in Sata. The results by Vakkari and Pöntinen (2015) on the association of an enriched results list to successful book discoveries support the findings of the Study V.

Overall, it seems that implementing diverse features to display the content of fiction books in library catalogs is associated to the book selection process in a positive way. The possible connection between the content description and the book selection process is interesting. If the enriched content description engages the user in the book search and selecting process more compared to the query-based book search in a traditional catalog, the enriched features could have an impact on enhancing the user experience in fiction book search in online library catalogs.

To conclude, even though differences occurred in users' search behaviors between catalogs, in general, the users of both catalogs were as successful in selecting interesting titles. In a traditional catalog a) a query is a necessary condition for browsing results, and b) returning to the search results from a book page is a necessary condition for proceeding to the next interesting item. Thus, in a traditional catalog, in both known-author and browsing tasks, effort needs be invested especially in well-designed queries, examining search results for clues of interesting topics, and exploring book pages in detail. A traditional catalog performs well for tasks with a known author, title (see Adkins & Bossaller, 2007), or topic (with advanced search). In an enriched catalog, the visual front page and the automatic recommendations enable the participants a) to browse without querying, and b) to proceed from item to item without the need to return to the search results. An enriched catalog performs well for open-ended browsing tasks that aim for serendipitous discovery.

Selection of a particular item. The interest criteria for novels as described by genuine fiction readers appeared to bear similarity to the relevance aspects identified by Koolen et al. (2015) and Reuter (2007). Even though previous research had identified familiarity to be an influential factor in fiction book selection, readers' literary knowledge, literary preferences and capacity to come up with authors and titles seem to have a remarkable role in the book selection process, particularly in browsing tasks. As the common selection strategies in public libraries (such as browsing the shelves of returned loans and new items) are not applicable in digital environments, readers are to discover alternative ideas for their book searches.

As in Reuter (2007), the participants of the Study V emphasized the role of a novel's title in the book selection. It would appear that the attributes for a good title are context related. In searches for truly interesting readings, the titles are carefully examined and clues to personal links to the subject of a novel are sought. Preliminary reading expectations are commonly produced based on title names in the search results. In searches for topical novels, the titles are examined in a quest for well-known classics about the topic instead of interestedness and engagement in the reading. Perhaps surprisingly, the readers in the Study V emphasized bibliographic information over the content in each task where a genuine reading interest was the point of departure for book selection. Only when the topic was pre-decided, the content description was emphasized over title and other bibliographic information. This suggests again that selecting topical novels resembles searching for non-fiction books and offers an interesting perspective to the discussion (e.g. Macgregor & McCulloch 2006; Spiteri 2009) on the role of professional metadata and user-generated content in library catalogs. Confirming the notions of Goodall (1989) and Saarinen and Vakkari (2013), it seems that in library catalogs, a combination of a user's previous knowledge and bibliographic information such as a title, trigger the preliminary interest and a curiosity to examine a novel in detail, whereas the content description, provided either by library professionals or other users, determines the final decision for selecting a book. The richer the metadata on item pages, the easier the decision for selecting a particular item is to make. The outcome of incomplete or missing metadata on an item page is commonly a book rejection, as the user fails in interpreting the possible reading experience. This is consistent with the findings of McKay et al. (2012).

The findings of the Study V indicate that a successful fiction book search process in a digital environment requires not only the selection of appealing items, but also the rejection of non-desirable items. Compared with book selection, rejection appears to be a quick and a straightforward process where a single attribute in bibliographic information or content description determines the novel to be non-interesting. The findings in Vakkari et al. (2014) are in line with the previous as assessing novels as non-interesting was found to require less effort compared with assessing them as interesting.

Quite the contrary, a book selection seems to be a multipart chain of search actions, during which the influence of literary knowledge, situational reading expectations and affective factors are associated to the outcome. In library catalogs, a successful book selection combines successful information retrieval, reader's search skills, a strong capacity to apply one's knowledge on preferred literary genres,

authors and topics, capacity to identify appealing items from a large amount of available readings, the interpretation of the expectations provoked by a particular title and system's support (such as a well-functioning search engine, complete and enriched metadata, a good recommender system). The challenge is to articulate one's idea of good reading as it is often unknown, especially at the beginning of the search. Compared with Ross's (2001) findings of the book selection process in physical environments, it seems to be even more challenging in library catalogs. *"Behind the eyes knowledge"* offers an effective starting point for the search but alone it is not sufficient for a successful book selection. The findings suggest that novels' interest criteria vary within successful selection processes. According to a search scenario, readers create expectations toward the reading material to be selected and apply the interest criteria alternately to meet these expectations. The interest criteria for novels identified in the Study V seem to be well applicable to various search contexts for recreational readings, and readers seem to be skilled in altering the emphasis between the interest criteria in a dynamic way according to the search scenario.

5.3   Fiction book search and casual-leisure scenarios

In this section, the findings of this thesis are reviewed in the light of previous knowledge on casual-leisure information behavior and the contribution of the present study for the research methods applied for examining casual leisure information behavior is discussed. Moreover, it is elaborated how the present findings could be related to the authentic search situations of recreational reading materials, and to the search for non-fiction information resources.

Elsweiler et al. (2011) have acknowledged the gap in our understanding for information behavior in non-work situations as including aspects such as what are the scenarios that motivate leisure-time information needs and what kind of factors can influence them. According to Elsweiler et al. (2011), the following factors

distinguish casual-leisure search scenarios from the standard models of information behavior: 1) the search tasks are often motivated by affective factors such as wanting to achieve a particular mood, 2) in casual-leisure situations the finding of information is often of secondary importance to the experience of finding, 3) in casual-leisure search the information needs are under-defined or absent, and 4) in casual-leisure tasks success is often not dependent on actually finding any information or results. In casual-leisure situations, there appears to be no surrounding task or activity, and casual-leisure behaviors can have a wide range of motivating factors affecting the search behaviors. Deviating from work-based scenarios, in casual-leisure situations the focus in the search seems to be the experience of searching, such as the engagement or enjoyment. A failure to resolve a particular need appears to have less importance in casual-leisure situations in comparison with work-based scenarios (Elsweiler et al. 2011). In the previous section, the fiction book search in a digital environment was presented as a model where the search context, the reader's background, the system related factors, the common search tactics and the factors influencing the selection of a particular item interconnect. Even though the fiction book search appears to begin with a desire for "just good books", the context of the search functions as the reference point against which the search process is evaluated. The ultimate target of the fiction search is the discovery of interesting readings, and searching without finding is experienced as frustrating by fiction readers. Therefore, in the light of the present findings, it would appear that the search for recreational readings in a digital environment does not unambiguously follow the aforementioned principles for casual-leisure situations. Instead, it seems that the situations for fiction search are positioned somewhere between casual-leisure and work-based scenarios.

It is true, that the fiction search is motivated by a range of factors and the motivation for fiction reading and one's literary preferences have a remarkable significance to the fiction book selection. Affective dimension is constantly involved in the fiction search process as creating expectations towards the novels to be selected and as influencing the selection and rejection of a particular item. The experience of finding seems to become apparent in an interaction with an enriched library catalog which engages the reader in the search process. However, simply passing of time and enjoying oneself does not seem to be essential in the fiction search process, at least in an interaction with a traditional library catalog. The issue of engagement and pastime might be a different one in platforms such as the LibraryThing, which aims at being both a social network and a library catalog, or Whichbook, which aims at serendipitous book discoveries based on the affective

dimension and reading experience provided by particular works of literature. It holds true that information needs in the fiction search are ill-defined. Nevertheless, they exist and the fiction search does not begin in a void. The Studies II, IV and V illustrated that at the beginning of a search scenario, a failure in discovering ideas for interesting reading based on literary knowledge appears to be a major obstacle in the book selection process. Like in Ooi & Liew (2011), it seems that in a natural setting readers tend to have an author or a title in mind when the search process for a fiction book is initiated. Thus, in authentic situations the overall need for the search may be undefined, but readers orientate themselves to the search process with a particular item or an author, a recommendation from an acquaintance, a literary awarded item or a book review from the social media. When it comes to the notion of success not dependent on actually finding any results, in the fiction book search it appears not to hold good. The conversations with the participants in this thesis revealed that during the experiment but also "in the wild", a failure to discover fiction books while in a quest for them is considered as an unsatisfactory outcome.

Thus, the fiction search seems to share similarities with both casual-leisure and work-based scenarios. Since the fiction book search process has not been previously studied from a casual-leisure perspective, this can be considered as new information. According to Elsweiler et al. (2011), the simulated work-tasks situations (Borlund 2003), the commonly applied approach in performing laboratory experiments in interactive information retrieval, would often be inappropriate for casual-leisure system evaluations. However, this thesis made use of the IIR evaluation model and applied the idea of simulated work tasks situations in the studying of the fiction book search process and evaluation of two online library catalogs. In general, this thesis showed that the principles of the IIR evaluation model seem to be well applicable for the studying of fiction book search. The simulated search scenarios were designed based on previous knowledge on the authentic situations where fiction search typically occurs, and both realism and researcher's control was included in the experimental design. The experiment was conducted with an authentic web service designed particularly for fiction retrieval as a test system, and the participants for the experiment were selected as representing authentic, real users with genuine fiction reading interest. The evaluation metrics used in the experiment may be seen as a combination of both traditional and alternative performance measures. *A task completion time* and *time taken to various search moves* are examples of traditional performance metrics in interactive information retrieval. However, in the present study the time based metrics were used typically to enlighten the user performance between various search scenarios and systems in order to produce basic information

on which search actions are the users devoting time in reality. Relevance judgments given by the participants' in Studies II-V were twofold:

> During the search process the participants made binary relevance judgments deciding whether the items encountered were interesting enough to be selected, or non-interesting and to be rejected.

> After selecting three novels in each search scenarios, the participants assessed the novels' interest level with a three-dimensional scale: a little interesting, a somewhat interesting and a very interesting.

In addition to applying the relevance judgments of the selected items, the thesis made use of qualitative data of interviews and conversations with the participants. As a result, a five-dimensional categorization of the influential factors in the fiction book selection was created and applied in analyzing user behavior in different search scenarios. The qualitative data analysis contributed to understanding about the user behavior more broadly than using exclusively traditional performance metrics and the relevance judgments would have done. Thus, in the light of this thesis, it could be considered as a recommendation for future studies exploring casual-leisure search situations to apply methodological triangulation to enhance the evaluation of user performance.

The search scenarios in Studies II-V were designed based on previous knowledge on the common search scenarios where fiction book search typically occurs. The known item search, the open ended browsing, the similarity based browsing and the search for vacation reading were commonly experienced as the natural ways of obtaining fiction books by the participants. Therefore, the findings considering the search behaviors of these search scenarios seem to be applicable to readers' natural behaviors, at least to navigational patterns and query techniques. Even so, in authentic search situations, fiction readers' interaction with library catalogs may remain to consist mostly as locating and reserving known items. However, the similarity based approach revealed in Study IV is likely to increase popularity as the users familiarize themselves with the functionality of recommender systems and the recommendations evolve in a more personalized direction. The obstacles in designing queries are to be avoided when similarity based browsing by proxy allows readers to discover similar reading experiences without the need to articulate the idea of the similitude in the form of search queries.

Already Waples (1940) noted that readers do not commonly read fictional books based on their topic. This thesis yielded that this holds true also in fiction book

selection: the topical search for novels was experienced as an unnatural way for the readers to approach the fiction book selections. When encountering a search scenario for topical novels, the participants altered their search behaviors as resembling the one for non-fiction information resources: detecting the topical relevance of the novels found and selecting books based on their topics instead of particular interestedness. Thus, the search behaviors for topical novels resembled the one for non-fiction information resources. It could be elaborated that in addition to topical search, also the other search scenarios could be applicable in the search for non-fiction information resources in digital environments. For example, in searching for research articles from databases, the search process is initiated with a composition of a query (Vibert et al. 2009). After receiving the preliminary search results, similarity based tactics could provide an effective alternative for the keyword searches. Tools such as "Find similar results" have emerged in scientific databases and increased popularity as they provide a more effortless means of novel discoveries instead of conducting keyword searches. Therefore, the navigational patterns in the similarity based tactic elicited in Study IV could be considered as having similarities to the search behaviors for non-fiction information resources. Likewise, the focused querying and topical search tactics could be practical also in the search situations other than fiction. However, the central distinguishing factor, the affective dimension, in the fiction book search process should be acknowledged. In fiction book selections, the context of the search is often emotion-bound, which affects the overall search process and the items to be selected. Also, in fiction book selections the evaluation of a particular item's capacity to provide a certain reading experience is of high importance. In search for non-fiction information resources, the document's topical relevance and the document's value in resolving a particular need or a task bears significance over personal and emotion-bound factors. It could therefore be assumed that the process of selecting a particular item was more complex in fiction book search in comparison with a search for other information resources. Therefore, the suitability of the model of the fiction book search (Figure 12) in studies examining non-fiction search scenarios should be questioned, as the model is a compound illustration of both personal, affective and system related factors.

The thesis investigated the scenarios that motivate the search for leisure-time readings and the factors which influence them. The thesis presented the fiction book search process to share similarities with both casual-leisure and work-based search scenarios. The ill-defined information needs, a lack of a particular task to be solved and the affective dimension as affecting the search process were found to be similar

to the casual-leisure scenarios. The outcome of the search as in finding information appears to be common with work-based scenarios. The question of an engagement in the search process and the importance of the experience of finding remains to be unclear. Studies II-V gave hints that the enriched content description, the various visual and social navigational tools and the automatic recommendations would engage the fiction searchers in the book selection process more than the query-based selection process in a traditional catalog. As the selection process turns into an amusing and engaging flow instead of designing one query after another, the reading materials selected might begin to fulfill readers' expectations often unknown or unconscious.

5.4   Implications for system design and practice

This thesis was one of the first studies focusing on the functionality of various enriched features in an online library catalog and their association with a fiction book search. In general, users' search success was found to be equal in both library catalogs. Then again, the findings showed a clear difference in the user behavior between a traditional and an enriched library catalog. It appears that the metadata enriched catalog supported users to find interesting readings more quickly, with less effort and with a variety of search approaches in comparison with the traditional library catalog. In the enriched catalog, especially in browsing scenarios, the readers applied a variety of search tactics, often selecting the novels without issuing search queries and following the social and visual features provided at the starting page of the catalog. The results of the Studies II-V suggest that fiction readers perceive the enhanced browsing features as a fruitful approach to fiction selections in library catalogs. The enhanced metadata in search results and book pages might affect the book selections to be more informed in the enriched catalog compared with the traditional catalog, as they are based on a detailed and versatile metadata at the former. Quite the contrary, in the traditional catalog, the users needed more time to issue search queries, to examine the search results and to investigate the book pages to determine a novel to be of interest to them in comparison with the enriched catalog. Also, the search tactic in the traditional catalog was naturally focused querying with an author, a title or a topic as the catalog did not provide other search

features than issuing queries, which was perceived problematic by the fiction readers in browsing scenarios .

Overall, the findings of this thesis gave hints that in authentic search situations, readers' interaction with traditional online library catalogs remains to be bound to the catalog's basic functions: at present, locating and reserving of known items in library catalogs appears to be more popular among library users in comparison with explorative browsing behavior. However, the interviews and conversations with the participants in this thesis revealed that the readers are a) browsing, b) aiming at serendipitous discoveries, and c) seeking the clues of possibly interesting readings in digital environments. Nevertheless, the platforms for the aforementioned behavior are often other than traditional online library catalogs. The social media, in particular Facebook and Twitter appear to be places where readers' are enjoying themselves while passing time and tips of interesting readings are encountered without engaging oneself in the active searching behavior. Then again, the Amazon and various Finnish online bookstores appear to be places where readers are actively browsing for interesting readings, often via automatic recommendations based on previous reading experiences. Therefore, it appears that at present, the use of traditional library catalogs consists mostly of reserving and locating the items discovered elsewhere.

The matter for the practice is in discussing should something to be done in order to alter the role of online library catalogs. In the year 2015, the public libraries' websites in Finland received almost 40 million visits while the respective figure for the actual visits in public libraries was 50 million (Finnish Public Libraries Statistics 2015). Should the library practice be satisfied with the fact that when it comes to recreational reading material, users have unquestionably accustomed themselves to using their public libraries' websites to reserve books and renew loans? Or should there be an ambition to engage the users even more to a new sort of browsing behavior, and by consequence make a versatile use of all the effort invested in developing online library catalogs? In the light of this thesis, it can be acknowledged that the fiction readers find public libraries' online catalogs as user-friendly while searching for recreational readings with the focused querying -tactic. It appears that the enriched features (such as automatic recommendations, users' reviews, tags, user's book piles, cover image browsing) supporting browsing behavior in online library catalogs are still found unfamiliar by some of the users. However, as users get accustomed to the new enriched features and as the recreational readings continue to be provided more and more in a digital form (e-books), it is likely that the role of online library catalogs changes to a more versatile direction.

The findings of this thesis showed that fiction readers' search behaviors in two rather different online library catalogs differed drastically, but resulted in an equally successful outcome. It appears that the users of online library catalogs value the well-functioning advanced search -option which supports the focused querying -tactic. However, the thesis showed that more diverse user behavior was detected when various enriched features were provided in the library catalog in comparison with a traditional, a query based catalog. Therefore, a few suggestions on system design can be made based on these findings. An appealing starting page in the enriched catalog Sampo turned out to be a major distinguishing factor between catalogs. Also, time devoted to exploring visual and social navigational tools on Sampo's front page seemed to be a favored search approach, especially in explorative search scenarios. Thus, providing an appealing starting point for the search that piques curiosity and encourages users to explore the catalog thoroughly would enhance traditional online catalogs. The results suggest that issuing queries other than for a known author or title was somewhat problematic for users. When users rely on their pre-knowledge, the books selected tend to be from well-known authors and genres. Providing multiple access points for the collection would support the selection of novels with desired themes or topics and enhance new and unusual book selections.

Querying is a fruitful and common search approach, especially in tasks where the author or title is already known. However, the high number of unsuccessful query reformulation intervals in the traditional catalog Sata hints that the searchers had a limited understanding on how to formulate effective queries. Features such as Did you mean...? or semantic autocompletion could support searchers to avoid zero hit situations. In Sampo, the missing advanced search option would have been useful for some tasks. Even so, in Sampo, simple queries produced rich results, as the ontology allowed for query expansion. Thus, the optimal interface would provide both a single search box and an advanced search option together with an ontologization of a thesaurus. When searching for a certain topic, the participants tended to associate query terms with different dimensions of the novels, e.g., theme and emotion. The indexing of novels with both objective and subjective index terms (Adkins & Bossaller 2007) would support topical and thematic search.

Furthermore, as noted by McKay et al. (2012) and Koolen et al. (2012), content presentation and metadata are essential factors in determining books' usefulness when searching for books in digital libraries. In this thesis, the enriched metadata in SERPs and book pages in Sampo might have supported the searchers to identify interesting topical novels more efficiently compared to Sata. In Sata, incomplete or missing metadata might have caused similar user behavior as in the work of McKay

et al. (2012). Displaying solely the author and the title of a novel might have encouraged searchers to open topically relevant book pages for closer examination, and to devote much time to identifying the novel's topical relevance. If no additional metadata was provided, the novel might have been rejected or assessed as a somewhat interesting. Thus, as an implication for system design it could be acknowledged that enriched interface elements (such as a snippet, and a cover image) in SERPs might support readers to identify topical novels more quickly compared to the traditional result list that provides solely titles and author names.

The notable role of bibliographic metadata in the book selection process confirmed the importance of traditional bibliographic indexing in library catalogs. The results suggest that in fiction searching, bibliographic information such as the number of pages or publication year partly substitute the touching of books and it is essential in selecting between interesting and non-interesting items. A book as one whole entity continues to be vital in the selection process also in library catalogs and fresh ways of displaying bibliographic information could be emphasized in fiction indexing. Moreover, like in Koolen et al. (2012), displaying the content of a novel with both professional metadata and user-generated content might support readers to assess novel's interest level more efficiently compared to providing solely professional metadata. Judging the relevance of a document with limited metadata information on a traditional catalog was found to be challenging for searchers (Blandford et al. 2007). The missing blurbs, cover images and book reviews in Sata might have affected the participants to reject novels that would have been of interest to them if enriched metadata had been available. Thus, providing a diverse range of metadata elements in title pages supports readers to judge the books. In addition, options for efficient browsing could enhance the finding of similarly interesting titles (Mckay et al. 2014). When browsing the visual and social features in the enriched catalog, the participants perceived it as challenging to pick items for a detailed examination without a specific scope to the books displayed. Thus, offering possibilities to narrow the amount of browsable items by publication year, a genre or media recommendation could enhance the browsing experience in library catalogs. On the front page, providing categories such as "This week's most viewed novels," "Novels recently in the media," "The new chick-lit novels of 2015" or "Famous books/awarded books from the past five years" could better simulate the browsing experience in a physical library.

As in Koolen et al. (2015), the results of this thesis revealed a great challenge in designing well-functioning recommender systems for fiction books. As the previous reading experiences were often the most influential factor in readers' book selections,

a recommender system capable of taking the previous experiences into consideration would be of great support in fiction retrieval. When recommending novels within the same genre, from the same author or novels covering similar topics, the automatic recommendations seem to function mostly well. However, the shortcoming is in recommending novels with a similar mood or a narrative style or creating recommendations based on a combination of various similarity criteria. The designing of a recommender system capable of combining unknown wishes, emotions and personal expectations of a fiction reader in book selection is a challenging issue in future studies. As the Finnish legislation restricts a recording of any loan information, it is an even greater challenge for the system design to add personalized recommendations in public libraries' online catalogs. Also, a user-friendly information architecture in online library catalogs should be understood. In Sampo, the automatic recommendations often did not catch participants' interest as they were positioned in the right-hand corner of the book page. In Sata, fuller book descriptions and book cover images were positioned in the right corner of the book page but were still infrequently used during the experiment. Thus, when developing systems for fiction retrieval, both the positioning and the visual appearance of the recommendations, descriptions, and cover images should be designed to catch the eye. Book descriptions should be available on the book page, not behind a click on a separate window.

5.5   Implications for future studies

Overall, for future studies on interactive information retrieval and casual leisure-situations, this thesis provided detailed information on users' browsing behavior and navigational patterns in versatile browsing scenarios in digital environments. The model of the fiction book search could be applied in studying the retrieval of other emotion-bound phenomena, such as music or movies. Also, the search tactics for fiction and the readers' interest criteria for fiction books might be applicable in other studies examining casual leisure behaviors. Future studies examining the navigational patterns and search tactics identified in the present study in non-fiction search scenarios could offer an interesting perspective and contribute to the comprehensive

understanding about the differences and similarities between a search for fiction and non-fiction.

Moreover, the platforms such as Whichbook, What should I read next and LibraryThing are interesting attempts to support readers' serendipitous book discoveries without a need for conducting search queries. The evaluation of user behavior and the user perceptions of the effectiveness of these tools in comparison with the findings of this thesis would enhance the overall understanding about fiction readers' search behaviors, which is essential as digital fiction collections continue to grow. Future studies on fiction retrieval could also make use of the versatile data collected in this thesis, which analyzed the data a) by identifying search moves, b) by measuring the frequency and duration of the identified search moves and c) comparing them between the catalogs. Also, three suggestive search tactics for fiction book search were elicited based on observing the recorded search sessions. However, different encoding methods could be applied to the data in order to produce more detailed information especially on users' navigational paths. This information could be used to build explicit path models of fiction readers' search behaviors. Moreover, as noted when discussing the study limitations (see section 3.4.5), the data of this thesis was biased towards highly educated women. The data could be complemented with a greater number of male participants in order to create more generazible results. Also, the division of participants in two different reader groups based on their reading preferences can be considered as a suggestive approach. With a larger number of participants, there could be a varied means to categorize fiction readers in to preference groups, which could enhance the understanding about varied search behaviors between different reader groups. As the fiction e-collections are emerging in library catalogs and digital libraries, there continues to be a lot of space for futures studies focusing on fiction reading taking place in digital environments and to the search behaviors for e-books in digital libraries.

6   CONCLUSIONS

The prime motivation for undertaking this study was to contribute to the comprehensive understanding about search behaviors for casual-leisure reading material and fiction readers' book selection process in online library catalogs. The study investigated fiction readers' book searching in an experimental setting, proceeding from an idea that different literary preferences have an association with the readers' book selection process in an interaction with an information system. The study illuminated various factors associated with search success when searching for novels in digital environments, which is essential as digital fiction collections continue to grow. The study investigated the scenarios that motivate the search for leisure-time readings and the factors which influence them. The study confirmed that known item search and browsing are the most used strategies for finding books in public libraries. The study enlightened readers' typical search approaches for books in various search scenarios and described the differences in user behavior between different library catalogs. The study presented the fiction book search process to share similarities with both casual-leisure and work-based search scenarios. For future studies on interactive information retrieval and casual leisure-situations, the present study provided detailed information on users' browsing behavior and navigational patterns in versatile browsing scenarios in digital environments.

Understanding of fiction readers' search behavior in online catalogs will be essential as fiction collections become wholly digital and continue to grow. The study has several main contributions to the comprehensive understanding of the fiction book search process in digital environments. Firstly, the study revealed that fiction readers' search behaviors in a traditional catalog and an enriched catalog differ but may result in equal success. Providing multiple appealing entry points for fiction collections together with an enriched results list may result in successful, serendipitous discoveries. Furthermore, providing visual and social navigational tools may trigger ideas for search and consequently support the discovery of interesting new material. For the second, the study contributed on demonstrating that different preference patterns for fiction books were clearly associated to readers' search behavior. For the third, the study presented three common search tactics for

fiction books in library catalogs: 1. Focused querying, 2. Topical browsing and 3. Similarity based tactic. The study demonstrated that the search scenario determined the use of the identified search tactics. Searchers' literary knowledge and previous reading experiences were associated to the selection of a suitable search approach in each search scenario: at all times the starting point for the search was either a known title or an author, an interesting topic or genre or a previously read well-known title. The new works of literature were assessed in relation to the earlier reading experiences. For the fourth, the study illuminated that browsing task type was clearly associated to readers' document viewing behavior. Browsing for topical novels resembled the retrieving of topically relevant non-fiction documents. The association of dwell time on book pages and document value was found to be non-linear: most effort was required to select somewhat interesting novels. When browsing for good books, the association of dwell time on book pages and document value was found to be linear: most effort in terms of dwell time was required to assess the most interesting cases. Therefore, if dwell time on book pages is considered as relevance feedback, it should be interpreted with understanding of the search scenario.

For the fifth, the study showed the basic characteristics of fiction book selection in various search scenarios in library catalogs. The study confirmed that the fiction readers' book selection process in library catalogs resembled that in physical libraries. Readers' prior literary knowledge and the novel's title had a major role in book selection. However, book selection in library catalogs was found to be a highly context-related process, where a combination of readers' search capacities, "behind the eyes" knowledge, personal and affective factors and a well-functioning interaction with the system being used resulted in a successful book selection. Diverging from the previous, the rejection of a novel turned out to be a quick and simple process where a single (or missing) metadata element determined a book to be rejected.

Finally, the study contributed on revealing that the selection criteria for novels varied by the search scenarios. Readers applied the identified interest criteria in a flexible and multiphase way depending on the search scenario and phase. Differences in metadata elements and content description between the catalogs were associated to the book selection process. Depending on the catalog used, the emphasis on selecting interesting titles was on either searchers' literary knowledge and bibliographic information, or content description and engagement in reading. The possible connection between the rich content description and book selection process

suggested that users' engagement in selecting books in library catalogs could be enhanced by implementing rich metadata into search results and item pages.

A profound understanding of readers' book selection behavior in library catalogs is vital as fiction is increasingly accessed in digital environments. Designing well-functioning and user-friendly interfaces for fiction readers with varied search skills and literary interests requires both in-depth qualitative and large scale evaluative user studies. The results of this study offer a solid starting point for further examination of user behavior in online catalog fiction discovery and are a significant step toward a comprehensive understanding of search behaviors for casual-leisure reading material.

7 REFERENCES

Adkins, D. & Bossaller, J. E. (2007). Fiction access points across computer-mediated book information sources: a comparison of online bookstores, reader advisory databases, and public library catalogs. Library and Information Science Research 29, 3, 354-368.

Agosti, M., Fuhr, N., Toms, E. & Vakkari, P. (Eds) (2014), "Evaluation methodologies in information retrieval", Report from Dagstuhl seminar 13441, Dagstuhl Reports Vol. 3 No. 10, available at: http://drops.dagstuhl.de/opus/volltexte/2014/4433/ (accessed February 8, 2017).

Anfinnsen, S., Ghinea, G. & de Cesare, S. (2011). Web 2.0 and folksonomies in a library context. International Journal of Information Management, 31, 63–70.

Bainbridge, D., Cunningham, S. J. & Downie, J. S. (2004). Visual collaging of a music digital library. In Proceedings of International Society for Music Information Retrieval 2004. (Barcelona, Spain, October 10-14, 2004.) ISMIR 2004.

Bates, M. J. (1979). Information search tactics. Journal of the American Society for Information Science, 30, 205–214.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. Online Review 13, October 1989, 407-424.

Bates, M. J. (1990). Where should the person stop and the information search interface start? Information Processing and Management 26(5), 575-591.

Bawden, D. (2011). Encountering on the road to Serendip? Browsing in new information environments. In Innovations in information retrieval. Perspectives for Theory and Practice, A. Foster & P. Rafferty, Ed. London: Facet Publishing, 1-22.

Bergstrom, A. & Hoglund, L. (2014). A national survey of early adopters of e-book reading in Sweden. Information Research, 19(2). Available at: http://www.informationr.net/ir/19-2/paper621.html#.WUVo7uvyg9c

Blandford, A., Makri, S., Gow, J. et al. (2007). A library or just another information resource? A case study of users' mental models of traditional and digital libraries. Journal of the American Society for Information Science and Technology 58(3), 433-445.

Borlund, P. (2003). The concept of relevance in IR. Journal of the American Society for Information Science and Technology 54(10), 913-925.

Buchanan, G. & McKay, D. (2011). In the bookshop: Examining popular search strategies. Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, 11, 269–278.

Caplinger, V. A. (2013). In the eye of the beholder: readers' advisory from a cataloging perspective. Reference & User Services Quarterly, 52(4), 287-290.

Carter, R. & Lubinsky, J. (2016). Rehabilitation Research. Priciples and actions. 5th ed. St.Louis, Missouri: Elsevier.

Case, D. (2007). Looking for information. A survey of research on information needs, seeking and behavior. 2nd ed. Academic Press.

Cooper, W. S. (1971). A definition of relevance for information retrieval. Information Storage and Retrieval 7(1), 19-37.

Cunningham, S. & Bennett, E. (2008). Understanding collection understanding with collage. In Proceedings of International Conference on Asian Digital Libraries 2008. (Bali, Indonesia, December 2-5, 2008.) ICADL 2008.

Denzin, N. (2006). Sociological Methods: A Sourcebook. 5th ed. London: Transaction Publishers.

Eden, B. (2002). Using metadata to build an enriched library catalog. Library Technology Reports, 38, 65–70.

Ekvall, I-L. & Larsson, S. (1997). Edvin – a search system for fiction based on the experience of users' needs. Information Services & Use 17(2/3), 81-85.

Elsweiler, D., Mandl, S. & Lunn, B. K. (2010). Understanding casual-leisure information needs: a diary study in the context of television viewing. In Proceedings of IIiX'10, ACM, New Brunswick, New Jersey, USA, 25-34.

Elsweiler, D., Wilson, X. & Kirkegaard Lunn, B. (2011). Understanding casual-leisure information behaviour. In A. Spink & J. Heinström (ed.), New Directions in Information Behaviour (pp. 211-241). Library and Information Science, Vol. 1. Emerald Group Publishing Limited.

Elsweiler, D., Wilson, M. L. & Harvey, M. (2012). Searching for fun: Casual-leisure search. ECIR 2012 Workshop Proposals.

Elsweiler, D. & Harvey, M. (2015). Engaging and maintaining a sense of being informed: Understanding the tasks motivating twitter search. Journal of the Association for Information Science and Technology, 66, 264–281.

Emanuel, J. (2011). Usability of the VuFind next-generation online catalog. Information Technology and Libraries, 30(1), 44-53.

Eskola, K. (1979). Suomalaiset kirjanlukijoina. Helsinki: Tammi.

Fidel, R. (1985). Moves in online searching. Online Review, 9, 61–74.

Finnish Public Libraries Statistics. (2015). Received from http://tilastot.kirjastot.fi/index.php?orgs=1&years=2015#results

Gerrig, R. J. (1993). Experiencing narrative worlds. New Haven: Yale Univesity Press.

Gold, J. (1990). Read for your life: literature as a life support system. Ontario: Fitzhenry & Whiteside.

Goodall, D. (1989). Browsing in the public libraries. LISU Occasional paper No 1. Loughborough: Library and Information Statistics Unit.

Griswold, W., McDonnell, T. & Wright, N. (2005). Reading and the reading class in the twenty-first century. Annual Review of Sociology, 31 (Aug. 2005), 127-141.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. Public Opinion Quarterly, 70(5), 646-675.

Guthrie, J. T. & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.) Reading research handbook. Mahwah, NJ: Erlbaum, 403-422.

Gwizdka, J. (2014). Characterizing relevance with eye-tracking measures. In Proceedings of IIiX'14. ACM New York, NY, 58-67.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). Multivariate data analysis. Upper Saddle River, NJ: Prentice-Hall.

Hassan-Montero, Y. & Herrero-Solana, V. (2006). Improving tag-clouds as visual information retrieval interfaces. Proceedings of the 1 International Conference on Multidisciplinary Information Sciences and Technologies, October 25–28, 2006, Merida, Spain. Retrieved from http://www.yusef.es/improving_tagclouds.pdf.

Hessel, H. & Fransen, J. (2012). Resource discovery: Comparative survey results on two catalog interfaces. Information Technology and Libraries, 31, 21–44.

Hinze, A., McKay, D., Vanderschantz, N., Timpany, C. & Cunningham, S. J. (2012). Book selection behavior in the physical library: implications for ebooks collections. In Proceedings of Joint Conference on Digital Libraries 2012. (Washington DC, USA, June 10-13, 2011). JCDL 2012. ACM, New York, NY.

Horsman, J. (1990). Something in my mind besides the everyday: women and literacy. Toronto: Women's Press.

Huysmans, F. & Hillebrink, C. (2008). The future of the Dutch public library: Ten years on. The Hague: SCP.

Hypén, K. & Mäkelä, E. (2011). An ideal model for an information system for fiction and its application: Kirjasampo and Semantic Web. Library Review, 60(4), 279 – 292.

Ingwersen, P., & Järvelin, K. (2005). The Turn: Integration of information seeking and retrieval in context. Dordrecht, The Netherlands: Springer.

Iyer, H. & Bungo, L. (2011). An examination of semantic relationships between professionally assigned metadata and user-generated tags for popular literature in complementary and alternative medicine. Information Research,16(3). Available at: http://www.informationr.net/ir/16-3/paper482.html

Jennings, B. & Sear, L. (1986). How readers select fiction: a survey in Kent. Public Library Journal 1(4), 43-47.

Jiang, T. (2013). An exploratory study on social library system users' information seeking modes. Journal of Documentation, 69(1), 6-26.

Koolen, M., Kamps, J. & Kazai, G. (2012). Social book search: comparing topical relevance judgements and book suggestions for evaluation. In Proceedings of CIKM'12. ACM New York, NY, 185-194.

Koolen, M., Bogers, T., van den Bosch, A. and Kamps, J. (2015), "Looking for books in social media: an analysis of complex search requests", Proceedings of ECIR'15, LNCS 9022, Springer, 184-196.

Kraayakamp, G. & Dijkstra, K. (1999). Preferences in leisure time book reading: a study on the social differentiation in book reading for the Netherlands. Poetics 26, 203-234.

Krashen, S. D. (1993) The power of reading: insights from the research. Englewood, NJ: Libraries Unlimited.

Laplante, A. & Downie, J. S. (2011). The utilitarian and hedonic outcomes of music information seeking in everyday life. Library & Information Science Research, 33, 202-210.

Lau, E. P. & Goh, D. H. L. (2006). In search of query patterns: A case study of a university OPAC. Information Processing and Management, 42, 1316–1329.

Liu, C., Gwizdka, J. & Liu, J. (2010). Helping identify when users find useful documents: examination of Query Reformulation Intervals. In Proceedings of IIiX'10. ACM New York, NY, 215-224

Liu, C. & Belkin, N. J. (2010). Personalizing information retrieval for multi-session tasks: the roles of task stage and task type. In Proceedings of SIGIR'10. ACM New York, NY, 26-33.

Liu, C., Belkin, N. J. & Cole, M. J. (2012). Personalization of search results using interaction behaviors in search sessions. In Proceedings of SIGIR'12. ACM New York, NY, 205-214.

Lukin, P. (1994). Where one can find good books? In Finnish. MA these. University of Tampere.

Macgregor, G. & McCulloch, E. (2006). Collaborative tagging as a knowledge organization and resource discovery tool. Library Review 55(5), 291-300.

Maker, R. (2008). Finding what you're looking for: A reader-centered approach to the classification of adult fiction in public libraries. Australian Library Journal, 57, 169–177.

Mar, R. A. & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. Perspectives on Psychological Science, 3(3), 173-192.

Marchionini, G. (2006). Exploratory search: from finding to understanding. Communications of the ACM 49(4), 41-46.

Maryl, M. (2008). The apology of popular fiction. Everyday uses of literature in Poland. In S. Zyngier et al (Eds.), Directions in Empirical Literary Studies (pp. 317-328). Amsterdam: John Benjamins Publishing Company.

McClellan, A. W. (1981). The reading dimension in effectiveness and service. Library Review 30, 77-86.

McKay, D. & Conyers, B. (2010). 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, 77-80.

McKay, D., Buchanan, G., Vanderschantz, C., Timpany, C., Cunningham, S. J. & Hinze, A. (2012). Judging a book by its cover: interface elements that affect reader selection of ebooks", Proceedings of OzCHI'12, ACM, New York, NY, 381-390.

McKay, D., Smith, W. & Chang, S. (2014). Lend me some sugar: borrowing rates of neighbouring books as evidence for browsing. In Proceedings of JCDL'14, 145-154.

Mehtonen, P. (2016). The library as a multidimensional space in the digital age. Information Research, 21(1). Available at: http://www.informationr.net/ir/21-1/memo/memo6.html#.WUVqL-vyg9c

Miesen, H. W. J. M. (2003). Predicting and explaining literary reading: an application of the theory of planned behavior. Poetics 31, 189-212.

Moulaison, H. L. (2008). OPAC queries at a medium-sized academic library: A transaction log analysis. Library Resources and Technical Services, 52, 230–237.

Moyer, J. (2007). Learning from leisure reading. A study of adult public library patrons. Reference and User Services Quarterly 46(4), 66-79.

Moyer, J. & Weech, T. L. (2005). The education of public librarians to serve leisure readers in the United States, Canada and Europe. New Library World, 106, 67–79.

National Endowment for the Arts (2012). How a nation engages with art: Highlights from the 2012 survey of public participation in the arts. National Endowment for the Arts Research Report no. 57, Washington, DC: National Endowment for the Arts. Retrieved from http://arts.gov/publications/highlights-from-2012-sppa.

Oksanen, S. & Vakkari, P. (2012). Emphasis on Examining Results in Fiction Searches Contributes to Finding Good Novels. In Proceedings of Joint Conference on Digital Libraries 2012. (Washington DC, USA, June 10-13, 2011). JCDL 2012. ACM, New York, NY.

Ooi, K. & Liew, C. L. (2011). Selecting fiction as part of everyday life information seeking. Journal of Documentation 67(5), 748-772.

O'Rourke, R. (1993). Unpopular readers: the case for genre fiction. Comedia working papers no. 7.

Patton, M. Q. (2002). Qualitative evaluation and research methods (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Pejtersen, A. M. & Austin, J. (1983). Fiction retrieval: experimental design and evaluation of a search system based on users' value criteria (Part 1). Journal of Documentation 39, (4), 230-246.

Pejtersen, A. M. & Austin, J. (1984). Fiction retrieval: experimental design and evaluation of a search system based on users' value criteria (Part 2). Journal of Documentation 40, (1), 25-35.

Pejtersen, A. M. (1989). The Bookhouse: Modelling user's needs and search strategies as a basis for system design. Risø report M-2794. Roskilde Risø National Laboratory. London: Comedia.

Peters, I. (2011). Folksonomies, social tagging and information retrieval. In Innovations in information retrieval. Perspectives for theory and practice. A. Foster & P. Rafferty Ed. London: Facet Publishing.

Pöntinen, J. & Vakkari, P. (2013). Selecting fiction in library catalogs: A gaze tracking study. Proceedings of the International Conference on Theory and Practice of Digital Libraries, 17, 72–83.

Radway, J. (1984). Reading the romance. Women, patriarchy and popular literature. Chapel Hill NC: North Carolina University Press.

Reuter, K. (2007). Assessing aesthetic relevance: children's book selection in a digital library. Journal of the American Society for Information Science and Technology 58(12), 1745-1763.

Rice, J. (1988). Serendipity and holism: the beauty of OPACs. Library Journal 113, 3, 38-41.

Richards, A. & Sen, B. (2013). An investigation into the viability of LibraryThing for promotional and user engagement purposes in libraries. Library Hi Tech, 31(3), 493-519.

Ross, C. S. (2001). Making choices: What readers say about choosing books to read for pleasure. The Acquisition Librarian 13(25), 5-21.

Ross, C. S., McKechnie, L. & Rothbauer, P. (2006). Reading matters. Westport, Conn: Libraries Unlimited.

Rothbauer, P. M. (2004). The internet in the reading accounts of lesbian and queer young women: failed searches and unsanctioned reading. Canadian Journal of Information and Library Science 28(4), 89-112.

Rowlands, I. & Nicholas, D. (2008). Understanding Information Behaviour: How Do Students and Faculty Find Books? Journal of Academic Librarianhsip, 34(1), 3-15.

Saarinen, K. & Vakkari, P. (2013). A sign of a good book: readers' means of accessing fiction in the public library. Journal of Documentation 69(5), 736-754.

Saarti, J. & Hypén, K. (2010). From thesaurus to ontology: The development of the Kaunokki Finnish fiction thesaurus. The Indexer, 28, 5–21.

Saracevic, T. (1996). Relevance reconsidered '96'. In Ingwersen, P. and Pors, N.O. (Eds), Proceedings of CoLIS2, Second International Conference on Conceptions of Library and Information Science: Integration in Perspective in Copenhagen, Royal School of Librarianship, Copenhagen, 201-218.

Saracevic, T. (2007). Relevance: a review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance. Journal of the American Society for Information Science and Technology, 58, 1915–1933.

Scales, A. M. & Rhee, O. (2001). Adult reading habits and patterns. Reading Psychology 22 (3), 175-203.

Schaller, R., Harvey, M. & Elsweiler, D. (2012). Out and about on museums night: Investigating mobile search behaviour for leisure events. In Proceedings of Searching4Fun Wksp, ECIR, 2012.

Scherf, E. M. (1990). Motive and goal as determinants in the treatment of texts: Some thoughts on a functional model of reception. Poetics 19(5-6), 487-504.

Shimray, S., Chennupati, K. & Chennupati, K. R. (2015). An Overview of Mobile Reading Habits. Journal of Library & Information Technology, 35(5), 344-354.

Slone, D. J. (2000). Encounters with the OPAC: on-line searching in public libraries. Journal of the American Society for Information Science 51(8), 757-773.

Spiller, D. (1980). The provision of fiction for public libraries. Journal of Librarianhsip and Information Science 12(4), 238-266.

Spiteri, L. F. (2009). The impact of social cataloguing sites on the construction of bibliographic records in the public library catalogue. Cataloging & Classification Quarterly 47(1), 52-73.

Statistics Finland 2009. Time use survey. Helsinki, Finland. Retrieved December 12, 2016 from: http://www.stat.fi/til/akay/tie_en.html

Stebbins, R. A. (1997). Casual leisure: a conceptual statement. Leisure Studies, 16(1), 17-25.

Stebbins, R. A. (2007). Serious leisure: a perspective for our time. New Brunswick, NJ: Transaction.

Stockmans, M. J. W. (2003). How heterogeneity in cultural tastes is captured by psychological factors: a study of reading fiction. Poetics 31, 423-439.

Tabachnick, B. G. & Fidell, L. S. (2013). Using multivariate statistics. 6th ed. Boston: Allyn and Bacon.

Tang, M. C., Sie, Y. J. & Ting, P. H. (2014). Evaluating books finding tools on social media: a case study of aNobii. Information Processing & Management, 50, 54–68.

Tang, M. C. (2016). The effect of online word of mouth on fantasy readers' preference stability: The moderating role of involvement and knowledge. Library & Information Science Research, 38, 292-300.

The Reading Agency (2013). "Library facts", available at: http://readingagency.org.uk/news/ library-facts004/ (accessed December 12, 2016).

Tepper, S. J. (1998). Why Do More Women Read Fiction? Center for Arts and Cultural Policy Studies, Princeton University. Paper Series No 6. Available at: https://www.princeton.edu/~artspol/workpap/WP06%20-%20Tepper.pdf

Tepper, S. J. (2000). Fiction reading in America: Explaining the gender gap. Poetics 27(4), 255-275.

Thudt, A., Hinrichs, U. & Carpendale, S. (2012). The Bohemian Bookshelf: Supporting serendipitous book discoveries through information visualization. In Proceedings of

the SIGCHI Conference on Human Factors in Computing Systems. (Austin, Texas, USA, May 05-10, 2012). CHI '12. ACM, New York, NY.

Timperley, P. & Spiller, D. (1999). The impact of non-fiction lending from public libraries. LISU Occasional paper No. 24. Loughborough: Library and Information Statistics Unit.

Toms, E. (2000). Serendipitous information retrieval. In DELOS workshop on "Information seeking, searching and querying in digital libraries", 17-20.

Usherwood, B. & Toyne, J. (2002). The value and impact of reading imaginative literature. Journal of Librarianship and Information Science 34(1), 33-41.

Vakkari, P. (1991). Social structure, book reading and the function of public libraries. In P. Kaegbein, (Ed.), Studies on research in reading and libraries. München & London: Saur, 259-281.

Vakkari P. & Serola, S. (2012). Perceived outcomes of public libraries. Library and Information Science Research 34(1), 37-44.

Vakkari, P., Luoma, A. & Pöntinen, J. (2014) . Book's interest grade and dwell time in metadata in selecting fiction. In Proceedings of IIiX'14. ACM New York, NY, 28-37.

Vakkari, P. & Pöntinen, J. (2015). Result list actions in fiction search: An eye-tracking study. Proceedings of the IEEE/ACM Joint Conference on Digital Libraries, 15, 7–16.

Van Riel, R., Fowler, O., & Downes, A. (2008). The reader-friendly library service. Newcastle upon Tyne, United Kingdom: Society of Chief Librarians.

Vibert, N., Ros, C., Le Bigot, L., Ramond, M., Gatefin, J. & Rouet, J. (2009). Effects of domain knowledge on reference search with the PubMed database: An experimental study. Journal Of The American Society For Information Science & Technology, 60(7), 1423-1447.

Ville´n-Rueda, L., Senso, J. A. & de Moya-Anego´n, F. (2007). The Use of OPAC in a Large Academic Library: A Transactional Log Analysis Study of Subject Searching. Journal of Academic Librarianship 33(3), 327-337.

Waller, V. (2010). Accessing the collection of a large public library: an analysis of OPAC use. Library and Information Science Research Electronic Journal, 20(1), 1-27.

Waples, D. (1940). What reading does to the people. Chicago UP: Chicago.

Weaver, M. (2007). Contextual metadata: faceted schemas in virtual library communities. Library Hi Tech 25(4), 579-594.

Wildemuth, B. M. (2004). The effects of domain knowledge on search tactic formulation. Journal of the American Society for Information Science & Technology, 55(3), 246–258.

Wilson, M. L. & Elsweiler, D. (2010). Casual-leisure search: the exploratory search scenarios that break our current models. HCIR'10. New Brunswick, NJ, 28-31.

Xie, I. & Joo, S. (2010). Transitions in search tactics during the web-based search process. Journal of the American Society for Information Science & Technology, 61, 2188–2205.

Yu, L. & O'Brien, A. (1996). Domain of adult librarianship. Advances in Librarianship, 20, 151–190.

Yu, L. & O'Brien, A. (1999). A practical typology of adult fiction borrowers based on their reading habits. Journal of Information Science 25(1), 35-49.

Zhang, Y. & Kudva, S. (2014). E-books versus print books: Readers' choices and preferences across contexts. Journal of the American Society for Information and Technology, 65(8), 1695-1706.

APPENDICES.

Appendix I. The Pre-Questionnaire

_____

Background variables

Question 1.

Gender

male

female

Question 2.   Year of birth   19____

Question 3.

Basic education

= primary school

= comprehensive school

= Matriculation examination

Question 4.

Vocational education

= none

= secondary modern school

= higher vocational diploma

= Master's degree

Experience in searching for fiction

Question 5.

How often do you use online library catalogues for fiction searching?

= at least once a week

= at least once a month

=once in 2-3 months

=more rarely

=I have never used an online library catalog for fiction searching

Question 6.

How often do you use Satakirjastot / BookSampo for fiction searching?

= at least once a week

= at least once a month

= once in 2-3 months

= more rarely

= I have never used Satakirjastot / BookSampo for fiction searching

Question 7.

Library attendance: How often do you visit a public library to borrow FICTION books?

= at least once a week

= twice a month

= once a month

= few times in six months

= once or twice a year

= never

Fiction reading

Question 8.

How many fiction books you have read in the last 12 months_____

Question 9.

How much do you usually enjoy reading fiction books?

= not at all

= a little bit

= I have some interest

= quite a bit

= very much

= I don't know

Question 10.

Name three of your favorite AUTHORS:

1. _____

2. _____

3. _____

Thank you for your answers!

Appendix II. Reading Preferences Questionnaire.

Please select how important the following aspects are when reading fiction books:

	Not important	A little important	Somewhat important	Very important
Fiction reading provides...				
A possibility for daydreaming	1	2	3	4
Entertainment	1	2	3	4
Artistic experiences	1	2	3	4
Escapism	1	2	3	4
Aesthetic pleasures	1	2	3	4
An opportunity for relaxation	1	2	3	4
A possibility to experience beauty	1	2	3	4
Widening of vocabulary and ability to express oneself	1	2	3	4
Intellectual stimulation	1	2	3	4
Widening the reader's view of the world	1	2	3	4
Insight into foreign culture	1	2	3	4
Practical solutions to everyday life	1	2	3	4
Understanding of the world	1	2	3	4
New perspectives	1	2	3	4
A novel offers / a novel is...				
A portrayal of a reality	1	2	3	4
A chance for identification	1	2	3	4
Unfamiliar subject	1	2	3	4
Entertaining	1	2	3	4
Rich and elaborate language	1	2	3	4
Gripping plot	1	2	3	4
Text that resonates with feelings	1	2	3	4
Easy to read	1	2	3	4
Surprising features of narration	1	2	3	4
Interesting characters	1	2	3	4
Thought-provoking	1	2	3	4
Challenge for its reader	1	2	3	4
Factual information about the world	1	2	3	4
Historical events	1	2	3	4
An original style	1	2	3	4
A novel has won an award	1	2	3	4

Appendix III. Coding scheme for dimensions.

Dimension	Theme	Definition	Example utterances
Familiarity	Known author or title	Mention of a specific known title or an author, based on a personal reading experience.	'Åsa Larsson, I know this author very well. The newest book is called The Secondly Deadly Sin. I haven't read that one yet".(SKe79) 'The Captain's Daughter is one that I have read". (KSp40) 'Of course I know this author, but I have never read anything from him". (KSk35)
	Known item from media and book awards	Mention of a specific known title or an author, based on a recommendation or a book review in media (newspaper, blog, television show, etc.) or literary awards.	'I attended the Turku International Book Fair and there was a discussion on this book there. I will choose it". (KSa1) 'Of course, I want to read Popula by Pirjo Hassinen which is nominated for a Finlandia Book Prize". (SKm64)
Bibliographic information	Author	Mention of the author of the book not known in advance	'Annikki Karikasniemi, I have never heard of her".(KSv46)
	Title	Mention of the title of the book not known in advance	'I choose Becoming Father and Daughter, solely based on the title". (SKk85)
	Cover	Mention of the book's cover	
	Format	Mention of the format of the book	'The cover image affected my selection, in the cover there was an image of an old poster". (KSl36)
	Publication date	Mention of the age of the book	'It should be a paperback as I will put it in my suitcase". (KSk11)
	Length	Mention of how long the book is	'I prefer books published recently. I think I would choose a book published in the 21st century". (SKc53) ' 391 pages, it determines that I will not choose this one, it is way too long for a holiday trip". (KSy48)
Content	Genre	Mention of the genre of the book	"I want horror stories because they are easy and entertaining". (SKa51)
	Plot & Topic	Mention of the plot, topic or events of the story, including interpretation, expectations or questions	'I guess this is about different cultures, there are topics such as joy and shame". (SKe79) 'I will take a look at that book called Faithful. This is about a life in an Orthodox monastery. I am a Christian but I am interested in other religions". (KSf6)
	Subject headings	Mention of the subject headings contained in the "Subject headings" field	'I will choose this because there are colonialism and cultural differences in the subject headings". (KSf6)

Engagement	Interest/Liking	Mention of general interest, like or dislike in the book	"This is absolutely brilliant, fascinating, an incredible topic". (KSp40) "I wonder would it be boring to read Baby Jane, as I am not particularly really interested, but a little interested anyway…". (KSl36)
	Mood	Mention of the mood of a particular book and mention of a wished mood for a book	"Definitely this one, this seems to be totally senseless". (KSq17) "I want to be entertained, I don't want to become anguished". (KSq17).
	Serendipity	Mention of an unfamiliar and serendipitous discovery	"It says Mika Waltari of France. Ooh, I can't believe this! Is this possible? What an amazing discovery!" (KSe5) "These three were found by chance, and actually I found more interesting titles that I expected". (SKm87)
Sociocultural	Social book sharing	Mention of friends, family members, teachers etc. who have recommended the book not known in advance	"My mother has this book at home, I have never read it. She has recommended this to me so I will choose it". (SKv96)
	Personal connection	Mention of personal connection to some aspect of the book	This says that the milieu of the book is Tampere and it is about students and universities. I am interested as I have studied and lived in Tampere for a year now". (KSc3)
	Intertextuality and similarity between various media formats	Mention of another book, television show or movie.	"This is 14 knots to Greenwich. The topics are universities, Great Britain, stepbrothers. It reminds me of the book called Brideshead Revisited, there could be similarities". (SKn88) "I will choose Berlin Poplars by Anne B. Ragde. I have watched the television show based on this novel". (KSl36)

ORIGINAL PUBLICATIONS

Readers' search strategies for accessing books in public libraries

Anna Mikkonen
University of Tampere
School of Information Sciences
Tel + 358-405301815

Anna.Mikkonen@uta.fi

Pertti Vakkari
University of Tampere
School of Information Sciences
Tel +358-505288354

Pertti.Vakkari@uta.fi

ABSTRACT

The aim of this study was to survey typical search tactics used by readers in public libraries. We analyzed whether demographic factors, the frequency of library visits, reading activity and the type of books read were associated with major search tactics. The data consisted of 1000 respondents of mail questionnaire collected in 2010 as a stratified random sample of the Finnish population aged 15 to 79 years. The most often used search tactics for accessing books in public libraries were known book or author search together with browsing search. The use of known item search strategy was associated with high educational level and the amount of non-fiction books read. The use of browsing search strategy was associated with the frequency of library visits, gender and the amount of fiction books read.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *search process*.

General Terms

Human Factors

Keywords

Public libraries, Fiction, Fiction retrieval, Search tactics

1. INTRODUCTION

Recreational reading is an important part of people's everyday life in many countries. The time use survey from Finland in 2009 indicates that reading continues to be one of the most popular leisure interests among Finnish population. For example, in 2009 75 % of people over the age of ten had read at least one book in the prior 12 months. 46 % had read books at least once a week and 20 % read books daily [32]. In U.S. the overall adult literary reading rate in 2008 was 50 % in the prior 12 months [23].

Leisure reading focuses on fiction reading. In 2010 80 % of respondents of a representative survey in Finland had read at least one fiction book in the prior 12 months [28]. In the U.S. the greater reading of fiction explains the new growth in adult literary readers. In 2008 nearly half (47 %) of all adults had read fiction

[23].

Borrowing books and leisure reading are the most popular activities among library users, and most of the books borrowed in the public libraries are fiction [16, 22, 28]. In 2010 fiction's proportion of all book loans was 67 % in public libraries in Finland [16]. The popularity of recreational reading of fiction has a long history. Already Berelson [4] reported that 60-65 % of the total loans in public libraries in the U.S. were fiction.

Public libraries continue to be a major source of fiction [16, 17, 19, 40]. However, in recent years little attention has been given on how readers search for books in public libraries. Spiller [30] found that 54 % of the fiction borrowed was found by known item search, where the reader knows the name of the author or the title of the novel in advance. The rest, 46 % of the fiction was found by browsing in the library. In Spiller's [30] study browsing was limited to browsing the library shelves. However, browsing refers also to searching behavior where one is going through book displays, returned loans' shelves and book shelves in attempt to identify interesting novels to read. Browsing and known item search have been identified as commonly used search tactics to fiction also by Goodall [6], Yu & O'Brien [40] and Ross [24].

Thus, leisure readers are a considerable subset of the population and leisure reading concentrates mostly on fiction. Fiction readers make up a large portion of the public library clientele. They use various tactics when searching for books to read for pleasure in public libraries. However, little is known about how common these tactics are in the adult population and do they vary between reader groups [25]. Most of the previous studies have used convenience sampling, whereas large scale representative studies on the distribution of search tactics and how they are associated with reading interest are lacking. This information would help to gain a better understanding of leisure readers' tactics in accessing fiction and to design more fitting fiction search tools for varying reader groups with different reading interests.

The aim of this study is to survey in a larger scale typical search tactics used by readers in public libraries. We analyze whether reading activity and the type of books read are associated with major search tactics and whether demographic factors such as age, gender and education play a role in search tactics used by readers.

2. RELATED RESEARCH

The previous research is presented here from three perspectives. First, we introduce characteristics associated with fiction reading. For the second, the major search tactics used for fiction searching in public libraries are presented. For the third, we offer an insight into the factors that are associated with the use of these tactics.

2.1 Reading fiction

Characteristics affecting the tendency to read fiction have been shown in several studies. Women are more likely to read fiction and more likely to pursue leisure reading as a hobby than men [7, 30, 13, 18, 24]. Tepper [34] found that after controlling education level women's likelihood to have read a book of fiction in the prior 12 months was over 2 times greater compared to men. Women are also more likely to be heavy readers than males and they tend to read a greater variety of books than men [24, 34]. It has also been shown [i.e. 7, 24) that recreational reading declines with age so that heavy readers are more likely to be younger rather than older.

In addition to gender and age, education has a major impact on leisure reading as highly educated people read more than people with lower education [32]. Education is also positively associated with the frequency of reading fiction [25]. Kraaykamp & Dijkstra [14] examined social differentiation in book reading preferences using data from a Dutch nationwide survey. Educational resources were important in explaining reading preferences as a higher educational level and the stimulation of literary reading in secondary school determined reading preferences for complex books. Similarly Griswold et al. [7] point out that reading is associated above all with education in a way that formal educational attainment is a major determinant of literacy proficiency.

In summary, pleasure reading and searching of fiction occurs in everyday life environments often as a serendipitous activity. Socioeconomic characteristics such as gender, age and educational level and social motives towards literary reading are shown to be associated with different reading interests of pleasure readers.

2.2 Fiction searching in public libraries

It has been shown [36] that searching for non-fiction in public libraries is characterized by somewhat different factors than fiction searching. Serendipity has a much smaller role in non-fiction searching compared to fiction searching. In a survey of 399 respondents only 6 % had entered the non-fiction section specifically to browse the shelves [36]. Previous studies [24, 30, 40] have shown that fiction is found often by serendipity when browsing.

Previous studies have revealed some major search tactics applied to fiction search in public libraries. Search tactic is here defined as a practical book selection method that the reader uses to encounter desired books [29, 40]. Searching by author or skimming the returned loans are examples of search tactics. A search strategy, on the other hand, in this study, is defined as a group of search tactics which share similar basic characteristics. We distinguish between two strategies, *browsing* and *known item search* depending to what extent the searcher knows the object in advance.

Spiller [30] and Yu & O'Brien [40] have found that the methods readers apply in their fiction searching and selection can roughly be broken down into two types, which are browsing and searching for specific author or title. Compared to known item search browsing is a more complex search strategy where previous reading experiences, elements from books themselves, affective and personal factors and the informational needs of a person are present at the same time. [24]. Browsing as a method of book selection involves "looking around" in public libraries and skimming the shelves, book displays, returned loans and book exhibitions with the reader hoping to encounter interesting books [6].

According to Yu & O'Brien [40] searching for specific books relies as a major part on authors' names. Searching by author names was found to be as important as browsing [30]. In addition, combination of browsing and known item search was found to be very popular by Spiller [30] who found that almost 78 % of fiction searching was carried out with the use of the combination of these two search modes.

Pejtersen [21] has identified three search tactics which match to our research goals. Analytical search tactic is used when readers wish to access books about some topic like the French revolution. Search by analogy is generated when readers want something similar to a novel they have read. Browsing strategy is applied in situations when readers have only a vague idea of what they would like to read. Yu & O'Brien [40] add to these tactics category search, in which readers are looking for a particular category of books like genre and try to find books within that category. Rothbauer [26] and Ooi & Liew [20] have found that fiction readers rarely rely on librarians as sources of recommendations for interesting books to read.

2.3 The factors associated with use of search tactics

In a qualitative study of 16 fiction readers in a public library, Saarinen & Vakkari [27] found that the readers' typical search tactics varied by reading activity. Avid readers, characterized by constant reading of various types of novels, were using both known item search and browsing strategies, whereas occasional readers, to whom reading offered occasional relaxation and who read only periodically, were using known item search strategy only. Occasional readers typically knew in advance what they wished to borrow in the library, which is typical to known item search [27]. On the other hand Ooi & Liew [20] found when studying fiction selection of 12 adult fiction readers in a public library that participants tended to have a book title or author already in mind when they visited their public library, and fiction selection was thus often based on a pre-determined decision. Instead of browsing in public libraries, participants sought actively recommendations for interesting books outside libraries, where trusted family members and friends, book club discussions, mass media and book reviews played a major role [20] These findings are consistent to *alerting sources* and *cost in time and money* as presented in Ross's [24] model of fiction readers' book selection.

Browsing of fiction in Ooi & Liew [20] was usually focused on the "New Books" display and the "Book returns" section. Similarly Saarinen & Vakkari [27] found that the shelf for returned books was a common place for avid readers to begin their search. When searching fiction by browsing, readers' literary competence plays a major role: less competent readers' natural source of information is the book itself while more competent readers may rely on their knowledge on authors, titles or genres acquired through constant scanning of their everyday environments [24, 27].

Known item search strategy requires readers to know in advance the author's name or the title of the book. The role of the mass media, including book reviews in print, the Internet and social media and radio, is also important [20, 24]. Known item search is

a search strategy which is most supported by current library catalogs in public libraries. Readers perceive it easy to check the availability and location of a certain novel in the library catalog [27, 29].

Search tactics have also been studied from a psychological perspective [10]. Heinström [10] explored students' personality traits and their association to preferred search tactics. An energetic personality, high motivation, and positive emotionality were shown to enhance likelihood for incidental acquisition and a search style where a topic is spontaneously explored. In contrary, low motivation, stress, and insecurity reduced receptivity.

In all, while browsing and known item search are the major search strategies by fiction readers in public libraries, we have little information on the factors associated with the variation of these commonly used search strategies in different reading groups. In system design information on search strategies in different reader groups is vital in order to develop new tools that better support readers with different reading interests.

3. RESEARCH DESIGN

The aim of this study is to survey the popularity of the major search tactics of fiction books in public libraries. The specific research questions are as follows:

- How common are the major search tactics for finding fiction in public libraries?

- Are demographic factors associated with the search tactics used?

- Are the number and the type of books read associated with the use of search tactics?

3.1 Data

The data were collected in 2010 as a stratified random sample of the Finnish population aged 15 to 79 years by mail questionnaire for studying the outcomes of public libraries [38]. The number of respondents was 1000. Compared to the overall Finnish population aged 15 to 79 years, the sample was well represented by age, marital status, and geographic region [31] but it was biased toward female and highly educated [see 38]. Studies [11, 25, 37] indicate that females are more frequent library users compared to males, and highly educated people are more frequent library users compared to less educated people. By implication, the sample is biased towards more active library users. The greater response rate of those interested in the phenomenon observed is a common feature in surveys [8]. The bias in the sample implies that we have to control the effect of both gender and educational level in the results. More information of the sample, see Vakkari & Serola [38].

The questionnaire included questions concerning search tactics for accessing books in public libraries, respondents' reading activity and the type of literature read in the prior 12 months. The questionnaire contained also variables measuring demographic factors age, gender and educational level. These variables were chosen in the study as it has been shown that age, gender and educational level are associated to pleasure readers' reading activity and reading preferences [7, 13, 18, 24, 25, 30, 32, 34].

The major search tactics in public libraries were derived from the previous literature [6, 21, 30, 40] which were known title or

author search, browsing on the shelves, browsing on the shelves of returned books, catalogue use, asking librarians and browsing book displays. These tactics were represented as follows in the questionnaire: *I know the name of the book or the author, I browse on the shelves, I skim the returned loans, I use the library catalog, I ask a librarian and I browse book displays.*

The respondents were asked to indicate how often they used each of the tactics for searching for books in public libraries in an ordinal scale ranging from 1 to 4, where 4 was "often", 3 was "sometimes", 2 was "seldom" and 1 was "never". An option "I can't tell" was also given.

The measure used indicates search tactics used for finding books in general in public libraries. Thus the respondents were not asked how often they used these tactics for finding solely fiction. However, as the majority of books borrowed in public libraries are fiction, the measure reflects mostly tactics for finding fiction. Moreover, as these six search tactics were derived from the previous literature with focus on fiction searching, they mostly reflect tactics applied for fiction searching.

As an indication of reading activity the respondents were asked to mark the number of books read in the last 12 months, excluding textbooks read for studying. The respondents were also asked to indicate how many of these books were non-fiction books. With this information new variables, the number of fiction books read in the last 12 months and the percentage of fiction books of the total amount of books read in the last 12 months, were constructed. Demographic variables included in our study were gender, age and the level of education. For more information about these variables see Vakkari & Serola [38].

The educational level of the respondents was divided into three categories, referred to here as low, middle, and high. Low is primary school education which consists of eight years of studies. As the previous comprehensive school education in Finland prior to school reform in the 1970s, only the older population has studied under this system. Middle level education is comprehensive school education, which is general knowledge education. This compulsory education starts in the year of the child's seventh birthday and is finished when the syllabus of comprehensive school education has been completed, 9 years later. High education refers to education consisting of at least Matriculation examination and international matriculation examinations (IB, Gymnasieexamen), which is three to four years of studies after the comprehensive school.

4. ANALYSIS

A major aim of this study was to explore how common are the search tactics used for fiction searching in public libraries. The data were analyzed by using one-way analyses of variance (ANOVA) to determine whether age, gender, education and the number and the proportion of fiction books read was related to search tactics used. For the analysis we calculated the proportion of those respondents who had used a particular tactic at least sometimes. We included in the analysis only those respondents who had read at least one book in the previous 12 months. This was done as it is not likely that those not reading books would search for books in a library.

In order to analyze more in detail associations between search tactics and reading activity linear regression analysis was applied. For the economy of analysis the six search tactics observed were

reduced by factor analysis into two major *search strategies*. Factor analysis is a statistical data-reduction technique used to explain variability among observed variables in terms of fewer unobserved variables called factors. The loading indicates the proportion of variance a factor is able to explain of the total variance of a variable. Loadings vary between +1 and -1. The closer the loading is to one, the higher the dependence of the particular variable is from the factor [33].

The eigenvalue indicates the degree to which a factor is able to explain the total variance of the variables included in the analysis. If a factor has a low eigenvalue, then it is contributing little to the explanation of variances in the variables and may be ignored as redundant with more important factors [33]. The lower default value in SPSS for forming a factor is 1. The number of factors can be selected for interpretation either mechanically based on eigenvalue, or theoretically based on a theory or theoretically justified interpretation of the factors.

Linear regression analysis was used to analyze and model the two strategies produced by the factor analysis and independent variables. The latter variables were as follows:

- the number of fiction books read in the last 12 months

- the number of non-fiction books read in the last 12 months

- the percentage of fiction books of all books read in the last 12 months

- the frequency of public library visits

- respondents' gender, age and educational level.

When constructing regression models the method Enter was used to add independent variables into the models. Enter is a method where the researcher controls the entry of variables and the advancement of the regression process. In the Enter method the variables within the model are either predicted by research hypotheses or selected based on previous literature [33]. We chose the variables based on literature in an attempt to produce a model which would yield information to understand which among the independent variables are related to the dependent variable.

5. FINDINGS
We analyze first how various search tactics are associated with demographics and reading interest. After that we build regression models for predicting the use of major search strategies.

5.1 The use of major search tactics
To determine whether demographic factors, reading activity and the proportion of fiction read were associated with use of different search tactics, a one-way analysis of variance was conducted. As it is shown in Table 1, the search tactics where the reader knows the name of the book or the author was the most popular, followed by various browsing tactics, while using a library catalog was the least popular search tactic.

For the one-way analysis of variance we calculated the proportion of those respondents who had used a particular search tactic at least sometimes. An analysis of variance (ANOVA) yielded some significant differences in the use of the search tactics between different demographic and reader groups (Table 2).

Table 1. The use of different search tactics by readers (N=877).

Search tactics	% often	% sometimes	% seldom	% never
Known book or author	57	27	9	6
Browsing on the shelves	29	38	20	13
Skimming the returned loans	27	26	22	25
Library catalog	10	20	30	40
Asking a librarian	11	33	33	23
Browsing book displays	16	38	26	20

ANOVA indicated some significant difference in the use of the search tactics observed. Known item search was used more by those with high education and by those who read fiction compared to non-fiction readers. The respondents who had read at least 5 books in the last 12 months used known item search more often compared to respondents who had read less.

The tendency to browse on the shelves was more common among females than males. The readers whose proportion of fiction was at least 50 % browsed on the shelves more often than those with a proportion of less than 50 %. Similarly, the tendency to browse was associated with the number of books read as those who had read at least 5 books in the last 12 months browsed on the shelves more often than those who had read less. Browsing on the shelves was most common by those with mid-level education compared to other educational groups.

In the other browsing mode, skimming the returned loans, the associations were very similar. The tendency to skim the returned loans was more common with females and those who had read more fiction and more books on the whole. These differences in the use of browsing on the shelves and skimming the returned loans between genders and between different reader groups might be due to the greater reading activity of women [7, 13, 18, 24, 30, 34]. Also, according to Saarinen & Vakkari [27] avid readers, whose literacy competency is great, seem to browse more often than occasional readers.

The library catalog was used at least sometimes only by 30 % of the respondents, which makes it the least popular search tactic. In ANOVA significant differences were found only between females and males, as the tendency to search by using a library catalog was more common with males. Age and reading preferences did not have a significant effect on using the library catalog. Asking a librarian differentiated educational groups. Respondents with mid-level education had asked for help from a librarian most commonly. Significant differences were also found between age groups, as respondents over 45 years asked for help from a librarian more often than younger respondents.

Finally, differences in browsing book displays were in part in line with results of browsing on the shelves and the returned loans tactics. Females were browsing book displays more often than males and respondents who had read more than ten books were browsing book displays more often than respondents who had read less.

Table 2. A one-way analysis of variance of six search tactics.

Search tactic	Gender	Age	Education	% of fiction	# of books read/y
Known book or author	n.s.	n.s.	p < .05 high > low	p <.05 fiction 0,1- 100 % > fiction 0 %	p <.05 5+ b/y > 1-4 books/y
Browsing on the shelves	p <.001 females > males	n.s.	p < .05 middle > low, high	p <.05 fiction 50 + %> fiction - 50 %	p <.001 5+ b/y > 1-4 books/y
Skimming the returned loans	p<.05 females > males	n.s.	n.s.	p <.05 fiction 50 + %> fiction - 50 %	p <.001 5+ b/y > 1-4 books/y
Library catalog	p <.05 males > females	n.s.	n.s.	n.s.	n.s.
Asking a librarian	n.s.	p < .05 age 45 + > age - 45	p < .05 middle >low, high	n.s.	n.s.
Browsing book displays	p <.05 females > males	n.s.	n.s.	n.s.	p <.001 5+ b/y > 1-4 books/y

Note: n.s. = non significant

In summary, there were some significant differences in the use of major search tactics between different demographic and reader groups. Browsing on the shelves, skimming the returned loans and browsing book displays are search tactics which share similar basic characteristics. Differences in use of these three *browsing* tactics were significant between females and males and between different reader groups. The larger amount of books read a year, the more often the three browsing tactics were used. The proportion of fiction read was associated with the use of browsing the shelves and book displays: the larger the proportion of fiction, the more often browsing the shelves or book displays occurred. Interestingly the proportion of fiction was also associated with known book or author search. If respondents had read only non-fiction books, they were using known item search less often than those who had read at least one fiction book. Also interestingly, respondents' age differentiated significantly only asking for help from a librarian search tactic. Respondents over 45 years old were asking help from a librarian more often than younger respondents.

5.2 Search strategies

In order to examine more closely the associations of demographic factors, reading activity, the proportion of fiction read and the frequency of public library visits to the use of six search tactics subsequent multivariate analyses were conducted. For the economy of analysis a factor analyses was conducted in order to reduce the number of search tactics variables.

In factor analysis the extraction method applied was principal components and the criterion applied was Eigenvalues greater than 1. To obtain a solution as comprehensible and easy to interpret as possible, a rotation method direct oblimin was applied. The initial principal component analysis extracted two factors with an Eigenvalue larger than one, explaining a total of 62 percent of the variance. The proportion of the variance explained by the two-factor solution exceeded the recommended criterion of 60 % [9]. All items fitted well in the solution: all communalities exceeded 0.40.

Direct oblimin rotation resulted in the following interpretation of factors (Table 3): The first factor (F1: 43 %) was labeled Browsing, as the highest loadings were on the variables browsing on the shelves, skimming returned loans and browsing book displays. The second factor (F2: 19 %) was labeled Known item search, as the highest loadings were on the variables known book or author, library catalogue and asking for help from a librarian. The use of the library catalogue and asking for help from a librarian can be seen as situations where the reader has the name or the author of the book in mind and is asking for help to locate the item.

Table 3. The factors of search tactics' variables.

Search tactics	Factors		Communalities
	Browsing	Known item search	
Known book or author	.400	**.618**	.424
Browsing on the shelves	**.839**	.242	.704
Skimming returned loans	**.879**	.222	.776
Library catalog	.173	**.767**	.596
Asking a librarian	.206	**.733**	.535
Browsing book displays	**.836**	.385	.714
Eigenvalues	2.2602	1.147	
Cronbach's alpha	0.82	0.52	

Note: highest loading in bold.

Both Kaiser-Meyer-Olkin's test (KMO) for measuring the adequacy of the samples, and Bartlett's test for sphericity were performed to validate the use of factor analysis. The KMO measure was 0.76 (x2=1170.165, df=15, p = .000) which can be held as satisfactory. Bartlett's test indicated (p = .000) that the factor analysis was relevant [33].

For further analyses sum variables were formed on the basis of the factor analyses. For the browsing strategy browsing on the shelves, skimming returned loans and browsing book displays were selected. For the known item strategy known book or author, library catalog and asking a librarian were selected. The reliability of the new sum variables was tested by counting Cronbach's alpha. For the browsing strategy, alpha was 0.82, which is satisfactory. For the known item search strategy, alpha

was 0.52. This indicates a low reliability of the sum variable. Therefore we selected the individual variable "I know the name of the book or the author" for representing known item strategy because it was the most common search tactic of this type. Thus, browsing strategy consists of the sum of all variables having high loadings on the respective factor, and known item strategy is represented by the particular variable "known book or author".

5.3 Modeling the use of major search strategies

In order to analyze to what extent the two major search strategies were associated to reading activity, the proportion of fiction read, the frequency of library visits and demographic factors, multiple linear regressions were calculated. Explanatory variables were included into regression models in an Enter fashion where the variables within the model were selected based on previous literature and the researchers controlled the advancement of the regression process [33]. Explanatory variables were divided into three groups:

- demographic variables age, gender and educational level

- variables measuring reading preferences (the number of fiction books read in the last 12 months, the number of non-fiction books read in the last 12 months, the proportion of fiction books of all books read in the last 12 months)

- the frequency of respondents' public library visits.

Table 4 displays three linear regression models regarding browsing search strategy.

In Model A demographic characteristics were introduced. Research [7, 13, 18, 24, 30] indicates that tendency to read fiction is stronger for women than for men, and for highly educated than for people with lower education. Thus gender and educational level were included in the first model together with age. It was hypothesized that the amount of fiction read was associated with the use of browsing strategy. Together with demographic variables also variables measuring reading preferences were included in the model. Model A resulted in an adjusted R square of 0.148. The model explains about 15 % of the variation of browsing strategy. In this model, gender, the amount of fiction read and the amount of non-fiction read were significant predictors and the signs had the expected direction.

The aim of regression analysis is to build as simple and theoretically reasonable model which explains the adjusted variance as well as possible [9]. In order to do so, the variables which did not contribute significantly to the explanation of browsing strategy were removed from the model. In model A variables age, basic education and the proportion of fiction read were not significant and they were removed. This resulted in slightly improved adjusted R square of 0.174 (Model B). The model explains about 17 % of the variation of browsing strategy. As expected, gender (β=.237) had the most significant effect on use of the browsing strategy, as women were using it more often than men. The effect of the amount of fiction read was also significant (β=.223): the more fiction read, the more often browsing search strategy was used.

Table 4. Regression of browsing search strategy on demographic variables, reading activity, the type of books read and the frequency of library visits (N=845)

Explanatory variables	Standardized regression coefficients Browsing search strategy		
	Model A	Model B	Model C
Age	.027		**.097****
Basic education	.040		
Gender	**.204****	**.237****	**.221****
The number of fiction read	**.195****	**.223****	**.160****
The number of non-fiction read	**.155****	**.151****	
The proportion of fiction read	.067		
The frequency of library visits			**.310****
Adjusted variance explained	14.8 %	17.4 %	21.9 %
F	25.3, p < .001	65.9, p < .001	51.8, p < .001
* = p < .05 ** = p < .01 *** = p < .001			

In model B little unexpected positive effect of the number of non-fiction read (β=.151) in the use of browsing strategy was revealed. Timperley & Spiller [36] found that users searching for non-fiction books in public libraries rarely tend to browse. However, regression analysis here indicates that the more non-fiction books read, the more often browsing search strategy was used. Nevertheless, compared to the effect of the amount of fiction read, the effect of non-fiction read remains lower. The positive effect of the amount of non-fiction read to the use of browsing strategy is in part due to the slight collinearity between variables of the amount of fiction read and the amount of non-fiction read [33]. In this case collinearity between two variables means that most of those who read fiction read also non-fiction. Thus, fiction reading causes in part the association of non-fiction and browsing. However, it is plausible that readers search also non-fiction by browsing. It may be that category search is common for finding non-fiction books on a category like travel guides or biographies.

Finally, in model C respondents' frequency of public library visits was added into the model. This resulted in an improved adjusted R square of 0.219. The model explains about 22 % of the variation of browsing strategy. The frequency of library visits (β=.310) had the strongest effect on use of the browsing strategy, as those visiting the public library more frequently were using it more often compared to less frequent visitors. In model C gender (β=.221) and the amount of fiction read were also significant (β=.160) predictors, as women and more frequent fiction readers were using browsing strategy more often than others. Diverging from models A and B, in model C age was a significant (β=.097).Although the findings of the regression analyses agree to a certain extent with the past research, it must be noticed that the

predictors account for only 21.9 % of the variance. In model B it was shown that gender was associated to the use of the browsing strategy as females browsed more than males. In model C it was shown that the frequency of library visits had an association to browsing strategy as frequent library visitors browsed more than those visiting the public library more seldom. It appears that educational level did not contribute to the variation in browsing strategy.

Next we examine the use of the known item search strategy, which was represented by a particular search tactic known book or author search. Table 5 displays three linear regression models regarding known item search strategy.

In known item strategy our analysis proceeds similarly to the browsing search strategy. In model A age, gender, and education were included in the model as well as variables measuring reading preferences. This resulted in an adjusted R square of 0.076, which was low. The model explained only about 8 % of the variance of known item strategy. We removed the variables that did not contribute to the explanation of known item search strategy. In model B only the variables basic education, gender and the amount of non-fiction read were included. This resulted in a slight improvement of the model as the adjusted R square increased to 0.132. The model explains about 13 % of the variance of known item strategy.

Compared to the browsing strategy, the model (Model B) was quite different. In the use of known item search strategy the most significant predictor was basic education (β=.278), which did not contribute at all to the use of browsing search strategy. The higher the education, the more often known item search strategy is used. The second most significant explanatory variable was the amount of non-fiction read (β=.151). The more non-fiction read, the more often known item search was used. The third significant explanatory variable was gender (β=.127) as women were using known item search strategy more often than men. It should be noticed that the results account only for 13.2 % of the variance.

In model C respondents' frequency of public library visits was added into the model. This resulted in a weakened adjusted R square of 0.060. The model explains 6 % of the variation of known item strategy. Compared to the browsing strategy, the model C is very different. The most significant explanatory variable was again basic education (β=.196) as the known item search strategy was used more often by respondents' with higher education. The frequency of library visits (β=.100) and the amount of non-fiction read were also significant (β=.083) explanatory variables. However, the association between the frequency of respondents' library visits and the known item search strategy was not as strong as it was with the browsing strategy.

In summary, regression analyses yielded significant differences in use of two major search strategies. The use of browsing search strategy was best explained by two different regression models: first, gender, the amount of fiction read and the amount of non-fiction read accounted for 17.4 % of the variance; second, the frequency of library visits, gender, the amount of fiction read and age accounted for 31 % of the variance. The use of known item search strategy was best explained by gender, basic education and the amount of non-fiction read. With these variables the results account for 13.2 % of the variance.

Table 5. Regression of known item search strategy on demographic variables, reading activity, the type of books read and the frequency of library visits (N=852)

Explanatory variables	Standardized regression coefficients Known item search strategy		
	Model A	Model B	Model C
Age	-.020		
Basic education	**.203***	**.278***	**.196***
Gender	**.084***	**.127***	
The amount of fiction read	.045		
The amount of non-fiction read	**.127***	**.151***	**.083***
The proportion of fiction read	.048		
The frequency of library visits			**.100***
Adjusted variance explained	7.6 %	13.2 %	6.0 %
F	12.47, p < .001	47.52, p < .001	16.305, p < .001
* = p < .05 ** = p < .01 *** = p < .001			

6. DISCUSSION

Our study is the first nationwide survey of search tactics used for accessing books in public libraries. It confirms the findings of past research [30, 40] as known item search and browsing were the most popular search strategies used by readers. The tendency to browse was associated with gender as females browsed more often than males. However, the strongest predictors of browsing strategy were the frequency of public library visits and the number of fiction books read. Known item strategy was used most by respondents with high educational level. Together with education the most significant predictor of known item strategy was the amount of non-fiction read: it was used more often by those who read more non-fiction. Thus our study confirms that associations exists between fiction reading and tendency to browse, and non-fiction reading and tendency to use known item strategy.

The findings are next discussed with respect to three research questions. Implications for system design are presented in the end of the discussion.

6.1 The most common search tactics
Our first research question seeks to examine how common are the major search tactics for finding books in public libraries among

book readers. The known book or author search was used at least sometimes by 84 % of the respondents. The next most common search tactics were browsing on the shelves (67 %), browsing book displays (54 %) and skimming the returned loans (53 %). Our results specify and update the findings in Spiller [30] and Yu & O'Brien [40] about the popularity of fiction search tactics. Known item search is considerably more popular compared to various browsing tactics.

Asking for help from a librarian was used at least sometimes by 44 % of the respondents. In previous research it has been noticed that fiction readers rarely rely on librarians as sources of recommendations for interesting books to read [26]. In our study we did not ask what kind of situations readers typically consulted librarians. Thus we cannot distinguish for example between recommendations for books and asking for help to locate known item in the library. However, compared to known item search and browsing tactics, asking for a librarian's assistance was used more seldom.

30 % of the respondents were using catalogs at least sometimes making it the least popular search tactic. The use of library catalogs and libraries' online services show similar trends in the US where library websites were used by a third of Americans in 2010 [22]. In Finland recent statistics reveal that visits to public libraries' websites are increasing vastly, in 2010 there were over 57 million visits on public libraries websites [16]. Based on this information it could be assumed that the proportion of those using library catalogs and conducting searches outside libraries will grow in the future.

6.2 Search tactics in different groups
We were also interested to find whether demographic factors, the frequency of public library visits and reading preferences were associated with the use of six major search tactics. Among demographic factors the most significant differences occurred between males and females. Gender was associated with the use of browsing tactics, which were browsing on the shelves, skimming the returned loans and browsing book displays. Females browsed more compared to males. This association is most likely due to the females' greater tendency to read and especially to read fiction, as shown previously [7, 13, 18, 24, 25, 30, 34]. The browsing strategy has been shown to be popular among fiction readers [6, 30], which is confirmed by our results.

Age was associated only with asking a librarian. Respondents over 45 years asked for help from a librarian more often than younger respondents.

Educational level has been found to be a major factor affecting leisure time reading as people with high education read more than people with lower education [7, 14, 25]. We found out that respondents' basic education significantly predicted the use of known item strategy but not browsing strategy for finding books. Respondents with high education were using known book or author tactics more often than respondents with lower education.

The most interesting finding in the use of different search strategies occurred between different reader groups. The effect of the type of literature most read to the use of search strategies was significant both in one-way analysis of variance and regression models.

The variables affecting significantly the use of browsing strategy were the frequency of public library visits together with gender, the amount of fiction read and age. In another model gender and the amount of fiction and non-fiction read were the most significant explanatory variables. These associations can be interpreted by analyzing the mechanisms behind them. First, association can be found between the frequency of public library visits and the tendency to browse. Experienced library visitors are more likely to have accustomed to the library's ways of arranging and displaying books. Thus experienced library visitors have gained "behind the eyes knowledge" which helps them to use the browsing strategy. Experience is also likely to explain the fact that age was associated with the browsing so that older respondents were browsing more often than the younger. For the second, association can be found between gender (female) and tendency to read fiction. It has been shown previously that females read more fiction compared to males [7, 13, 18, 24, 30]. For the third, in previous research [24, 27] there had been signs that the use of browsing strategy would be associated with the amount of fiction books read. For example Saarinen & Vakkari [27] showed that avid fiction readers used browsing search strategy more often than occasional readers. Our study strongly supports these findings showing that significant association can be found between gender, the tendency to read fiction and the use of browsing strategy. The model including the frequency of library visits accounted for 21.9 % of the variance in browsing strategy. In the other model association between the amount of non-fiction read and the use of browsing strategy was revealed. Thus it is likely that readers also browse to some extent for finding non-fiction. This can be considered as new information, but it must be remembered that little is still known about searching for non-fiction in public libraries. In our study this association is likely to be in part due to the collinearity in regression models between independent variables the amount of fiction read and the amount of non-fiction read. Collinearity here indicates that most of those who read fiction read also non-fiction and the association of non-fiction and browsing is actually caused by fiction reading. The model accounted only for 17.4 % of the variance in browsing strategy.

Compared to browsing strategy the variables affecting significantly the use of known item strategy were respondents' gender together with their basic education and the amount of non-fiction read. Again we can interpret these findings by analyzing the mechanisms behind the revealed associations. First, the association between gender and the use of known item strategy occurred in a way that females were using known item strategy more often than males. This can be explained by the fact that females read more than males in general [7, 13, 18, 24, 30]. For the second, basic education was also a significant predictor of known item strategy as highly educated respondents were using it more often than respondents with lower education. Here the association is likely to be due to the fact that people with higher education read more non-fiction than people with lower education [7]. Thirdly, there had been some signs that in public libraries non-fiction is searched most by known item strategy [36]. The findings of our study are in line with previous research [36]. Thus our study indicates that the use of known item strategy is best explained jointly by gender, basic education and the number of non-fiction books read. The more non-fiction read, the more often the known item search strategy was used. The model accounted only for 13.2 % of the variance of known item strategy. The association between the frequency of public library visits and the use of known item search strategy was also revealed. However, compared to the browsing strategy the association was not as

strong, as the model including the frequency of library visits accounted only for 6 % of the variance of known item strategy. It is likely that the experienced library visitors are more accustomed to locate the books they have in mind in advance when they visit the public library than those visiting public libraries more seldom.

Our aim was to survey in a larger scale typical search tactics used by readers in public libraries. We analyzed whether reading activity and the type of books read were associated with major search tactics. Our study indicates that the most common search tactics used in public libraries, browsing on the shelves and known item search, have remained the same since from the 1980s [30]. The results of our study provide evidence that the use of different search tactics is associated above all with the frequency of public library visits, gender, educational level and the type of books read. In our study there was a strong association between reading of fiction and the use of browsing strategy. Respectively, association between reading of non-fiction and the use of known item strategy was clear. However, the results did not account for a large proportion of the variance of different search strategies. Thus, there is a lot of space for further studies identifying factors associated with the variation of search strategies for books in public libraries. For example the associations of factors such as the amount of time patrons have to spend, individual differences and readers' criteria for choosing books to the preferred search tactics remains unknown.

6.3 Implications for system design
Our study showed that readers searching for books in public libraries are using known item and browsing tactics most often. These search tactics varied between different reader groups. The use of the library catalog was the least popular search tactic, as only a third of the respondents were using catalogs at least sometimes. However, it is known that visits to public libraries websites are increasing every year [15, 16] and that public library users may have little time to use on browsing the physical library environments [20]. In system design information on the use of various search tactics between different reader and demographic groups could provide new approaches to the design of library systems and OPACs. Current library systems can be considered somewhat static as they do not adapt to meet the needs of different readers. For example readers wishing to locate known item in a library have different demands on library systems than those wishing to browse. Similarly current content description methods in library systems give only little support for readers wanting to find books similar to a known one. Also librarians are largely left to rely on their own reading, memory and associations when aiming to guide a reader towards a book similar to a known one [12].

The development of new search systems has mainly focused on retrieving non-fiction. In addition to Pejtersen's BOOK HOUSE model there are some studies [e.g. 1, 3, 5, 39] focusing on fiction retrieval in online catalogs in public libraries. Thudt et al. [35] have introduced a visual exploration tool, the Bohemian Bookshelf, which aims to support serendipitous book discoveries offering multiple visual access points to digital book collections. However, little is known how library systems support fiction search strategies other than known item search. As Bawden [2] points out, computer searching is good for finding specific information but still quite ill-suited to allow a browsing approach.

In an era where fiction retrieval occurs even more often in digital environments, the popularity of browsing should be taken into consideration when designing new library systems. New techniques supporting different browsing tactics as a part of fiction retrieval should be designed in order to meet the needs of fiction readers. For example more effort should be put on designing different ways of displaying the search results in OPACs. Faceted search discovery interfaces could help readers to continue searching without conducting new queries and different recommendations based on reader profiles could be provided for those wishing to read books from a certain genre. Similarly in known item search recommendations such as "more from the same author" displayed in different ways in OPACs could provide information for a more varied selection of novels to read. Larger scale studies analyzing user interactions in transaction logs in different reader groups could provide vital information on how users are really accessing fiction in current library catalogs and how successful are their searches. Based on this information different reader profiles could be provided and used in system design for example as a basis of recommendations.

7. CONCLUSION
This study examined the popularity of different search tactics used for finding books by readers in public libraries. Compared to previous research [18, 24, 40], this study was the first nationwide survey on the distribution of search tactics in different reader groups. Known author or book tactics was found to be the most popular, various browsing tactics following. Six particular search tactics were reduced to two major search strategies by factor analysis and regression models were calculated to predict the use of browsing and known item strategies. The use of browsing strategy was best predicted by the frequency of library visits, gender, the amount of fiction read and the amount of non-fiction read. The use of known item strategy was best predicted by gender, basic education and the number of non-fiction read. Although the results confirm and specify previous findings [24, 27, 40], the variance explained by regression models remained low. Thus many of the factors associated with the variation of search strategies for books in public libraries remain unknown. However, for system design the study offers evidence that known item search and browsing are the most used strategies for finding books in public libraries. To better respond the needs of readers with different reading, and thus searching preferences, the library systems should be designed to better adapt to individual needs of readers. Future explorations on displaying search results in multiple ways could prove to be useful for supporting better browsing tactics in online environments.

8. REFERENCES
[1] Adkins, D. & Bossaller, J. E. 2007. Fiction access points across computer-mediated book information sources: a comparison of online bookstores, reader advisory databases, and public library catalogs. *Libr Inform Sci Res* 29, 3 (Sep. 2007), 354-368. DOI = http://dx.doi.org/10.1016/j.lisr.2007.03.004.

[2] Bawden, D. 2011. Encountering on the road to Serendip? Browsing in new information environments. In *Innovations in information retrieval. Perspectives for Theory and Practice,* A. Foster & P. Rafferty, Ed. London: Facet Publishing, 1-22.

[3] Beghtol, C. 1994. *The Classification of Fiction: The Development of a System Based on Theoretical Principles Metuchen,* NJ: Scarecrow Press.

[4] Berelson, B. 1949. *The Library's public.* Columbia UP: New York.

[5] Carlyle, A. & Ranger, S. 2001. Facilitating retrieval of fiction works in online catalogs. In *Proceedings of the 12th ASIS SIG/CR Classification Research Workshop 2001* (Washington D.C., USA, November 04, 2001). Retrieved March 27, 2012 from: http://faculty.washington.edu/acarlyle/Papers/facilitating.pdf.

[6] Goodall, D. 1989. Browsing in the public libraries. *LISU Occasional paper* No 1. Loughborough: Library and Information Statistics Unit.

[7] Griswold, W., McDonnell, T. & Wright, N. 2005. Reading and the reading class in the twenty-first century. *Annu Rev Sociol*, 31 (Aug. 2005), 127-141. DOI= http://dx.doi.org/10.1146/annurev.soc.31.041304.122312.

[8] Groves, R.M. 2006. Nonresponse rates and nonresponse bias in household surveys. *Public Opin Quart*, 70, 5, 646-675. DOI= http://dx.doi.org/10.1093/poq/nfl033.

[9] Hair, J. F., Anderson, R. E., Tatham, R. L. & Black, W. C. 1998. *Multivariate data analysis.* Prentice-Hall, Inc: New Jersey.

[10] Heinström, J. 2006. Psychological factors behind incidental information acquisition. *Libr Inform Sci Res* 28, 4, 579-594. DOI = http://dx.doi.org/10.1016/j.lisr.2006.03.022.

[11] Huysmans, F., & Hillebrink, C. 2008. *The future of the Dutch public library: Ten years on.* The Hague: SCP.

[12] Hypén, K. & Impivaara, A. 2011. Read, describe and share! Building an interactive literary web service: an article about Kirjasampo. *Collection Building* 30, 1, 61-67. DOI= http://dx.doi.org/10.1108/01604951111105032.

[13] Jennings, B. & Sear, L. 1986. How readers select fiction: a survey in Kent. *Public Library Journal* 1, 4, 43-47.

[14] Kraayakamp, G. & Dijkstra, K. 1999. Preferences in leisure time book reading: a study on the social differentiation in book reading for the Netherlands. *Poetics* 26, 4 (May 1999), 203-234. DOI= http://dx.doi.org/10.1016/S0304-422X(99)00002-9.

[15] Library Statistics Finland 2009. Retrieved March 27, 2012 from: http://tilastot.kirjastot.fi/fi-fi/perustilastot.aspx?AreaKey=Y2009T1N1.

[16] Library Statistics Finland 2010. Retrieved March 27, 2012 from: http://tilastot.kirjastot.fi/fi-fi/perustilastot.aspx?AreaKey=Y2010T1N1#group6.

[17] Maker, R. 2008. Finding what you're looking for: a reader-centred approach to the classification of adult fiction in public libraries. *Aust Libr J* 57, 2 (May 2008), 169-177. Retrieved March 27, 2012 from: http://www.alia.org.au/publishing/alj/57/ALJ_Vol57_No2_2008.pdf

[18] Moyer, J. 2007. Learning from leisure reading. A study of adult public library patrons. *Ref User Serv Q* 46, 4 (Jun. 2007), 66-79.

[19] Moyer, J. & Weech, T. L. 2005. The education of public librarians to serve leisure readers in the United States, Canada and Europe. *New Library World* 106, 1/2, 67-79. DOI= http://dx.doi.org/10.1108/03074800510575366.

[20] Ooi, K. & Liew, C. L. 2011. Selecting fiction as part of everyday life information seeking. *J Doc* 67, 5, 748-772. DOI= http://dx.doi.org/10.1108/00220411111164655.

[21] Pejtersen, A. M. 1989. *The Bookhouse: Modelling user's needs and search strategies as a basis for system design.* Risø report M-2794. Roskilde Risø National Laboratory.

[22] Perceptions of libraries. 2010. Context and community. (2010) OCLC. Retrieved March 27, 2012 from: http://www.oclc.org/reports/2010perceptions/2010perceptions_all.pdf.

[23] Reading on the rise. A new chapter in the American literacy. (2009) National Endowment for the Arts'. Retrieved March 27, 2012 from: http://www.nea.gov/research/ReadingonRise.pdf.

[24] Ross, C.S. 2001. Making choices: What readers say about choosing books to read for pleasure. *The Acquisition Librarian* 13, 25, 5-21. DOI= http://dx.doi.org/10.1300/J101v13n25_02.

[25] Ross, C. S., McKechnie, L., & Rothbauer, P. 2006. *Reading matters.* Westport, Conn: Libraries Unlimited.

[26] Rothbauer, P. M. 2004. The internet in the reading accounts of lesbian and queer young women: failed searches and unsanctioned reading. *Can J Inform Lib Sci* 28, 4 (Dec. 2004), 89–112.

[27] Saarinen, K. & Vakkari, P. 2012. A sign of a good book: Readers' means of accessing fiction in the public library. Submitted for publication.

[28] Serola, S. & Vakkari, P. 2011. The role of public libraries in citizens' activities. A survey on the benefits of public libraries in everyday life. Publications of the Ministry of Education and Culture 21. Retrieved March 27, 2012 from: http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2011/liitteet/OKM21.pdf.

[29] Slone, D. J. 2000. Encounters with the OPAC: on-line searching in public libraries. *J Am Soc Inform Sci 51*, 8 (Apr. 2000), 757-773. DOI= http://dx.doi.org/10.1002/(SICI)1097-4571(2000)51:8<757::AID-ASI80>3.0.CO;2-T.

[30] Spiller, D. 1980. The provision of fiction for public libraries. *J Libr* 12, 4 (Oct. 1980), 238-266.

[31] Statistics Finland 2009. Population Structure. Helsinki, Finland. Retrieved March 27, 2012 from: http://www.stat.fi/til/vaerak/2009/index_en.html.

[32] Statistics Finland 2009. Time use survey. Helsinki, Finland. Retrieved March 28, 2012 from: http://www.stat.fi/til/akay/2009/akay_2009_2011-05-17_kat_004_fi.html.

[33] Tabachnick, B. G. & Fidell, L. S. 2001. *Using multivariate statistics.* 4th ed. Boston: Allyn and Bacon.

[34] Tepper, S. J. 2000. Fiction reading in America: Explaining the gender gap. *Poetics* 27, 4 (May 2000), 255-275. DOI= http://dx.doi.org/10.1016/S0304-422X(00)00003-6.

[35] Thudt, A., Hinrichs, U. & Carpendale, S. 2012. The Bohemian Bookshelf: Supporting serendipitous book discoveries through information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* (Austin, Texas, USA, May 05-10, 2012). CHI '12. ACM, New York, NY. Retrieved March 27, 2012 from: http://www.utahinrichs.de/uta/uploads/Publications/Publications/ThudtCHI2012.pdf.

[36] Timperley, P. & Spiller, D. 1999. The impact of non-fiction lending from public libraries. *LISU Occasional paper* No. 24. Loughborough: Library and Information Statistics Unit. Retrieved March 27, 2012 from: http://www.lboro.ac.uk/departments/ls/lisu/downloads/OP24.pdf.

[37] Vakkari, P. 1991. Social structure, book reading and the function of public libraries. In P. Kaegbein, (Ed.), *Studies on research in reading and libraries.* München & London: Saur, 259-281.

[38] Vakkari P. & Serola, S. 2012. Perceived outcomes of public libraries. *Libr Inform Sci Res* 34, 1 (Jan. 2012), 37-44. DOI= http://dx.doi.org/10.1016/j.lisr.2011.07.005.

[39] Van Riel, R. 1999. A book to match your mood: Forager information system helps readers select books based on psychological attributes. *Library Association Record* 101, 5 (May 1999), 292-293.

[40] Yu & O'Brien. 1996. Domain of adult librarianship. *Advances in librarianship* 20, 151-190.

Finding fiction: Search moves and success in two online catalogs

CrossMark

Anna Mikkonen *, Pertti Vakkari

School of Information Sciences, University of Tampere, FIN-33014, Tampere, Finland

A R T I C L E   I N F O

A B S T R A C T

Search moves for finding novels in five search tasks and two catalogs were analyzed. Search tasks reflected the following search tactics: known-author search, topical search, open-ended browsing, search by analogy, and searching without a query. The most used search moves in both catalogs across all tasks were querying, search results inspection, and book page examination. In a traditional catalog, more effort was needed in the form of queries, search moves, and opened book pages to gain equivalent average book scores when compared with an enriched catalog. In a traditional catalog, a typical search strategy for interesting titles seemed to involve issuing queries and considering suitable entry terms carefully, and devoting more attention to search results instead of book pages. In an enriched catalog, a common approach involved time devoted to exploring the catalog's enriched front page and multiple entry points together with attention to the enriched results list.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Leisure reading focuses on fiction (National Endowment for the Arts, 2012; Van Riel, Fowler, & Downes, 2008). Fiction is often searched for and found in public libraries (Maker, 2008; Moyer & Weech, 2005; The Reading Agency, 2013; Yu & O'Brien, 1996). In 2013, 44% of all book loans in public libraries in the United Kingdom were for fiction (The Reading Agency, 2013). Although it is known that leisure reading concentrates mostly on fiction, and fiction readers make up a large portion of the public library clientele, it is not known how readers are currently using library catalogs for fiction searching.

Research on search behavior in digital systems has mainly focused on retrieving non-fiction (e.g. Hessel & Fransen, 2012; Lau & Goh, 2006; Moulaison, 2008; Waller, 2010). Users' interaction with library catalogs in fiction retrieval is clearly an under-researched area, although research interest in fiction retrieval as a casual leisure activity is on the rise (Stebbins, 2007). For example, Buchanan and McKay (2011) examined search strategies used for finding books. Thudt, Hinrichs, and Carpendale (2012) have designed visualization techniques to support serendipitous book discovery in digital library collections. Oksanen and Vakkari (2012) have studied how enriched tools are used when searching for novels in a catalog. In addition, Pöntinen and Vakkari (2013) have examined how readers explore the fiction metadata in a gaze tracking study.

2. Problem statement

In recent years, public library online catalogs have been developed to meet the needs of diverse reader groups. Unique metadata that enrich the bibliographic description of a collection and enable users to browse and interact with the collection have become increasingly available in online library catalogs (Eden, 2002). The concept of an enriched library catalog (or a metadata-enriched library catalog) refers to various metadata elements and applications (such as book cover images, book descriptions, and virtual book shelves) that provide many different kinds of access points to library collections and allow users to read, add reviews, and generate content; these metadata supplement professionally created metadata. In spite of the development of enriched library catalogs and enhanced fiction content description, evaluation of these tools in fiction retrieval is scarce. Little is known about how fiction readers currently search for recreational reading material in enriched online catalogs as compared to traditional catalogs. A better understanding would inform the design of library catalogs for finding fiction. The study addresses the following research questions:

1. Which search moves are applied in different types of search scenarios for finding fiction?
2. Are there differences in search moves among different types of catalogs?

* Corresponding author.
  E-mail address: anna.mikkonen@uta.fi (A. Mikkonen).

3. Are there differences in search moves among different types of search tasks?

3. Literature review

3.1. Search moves and tactics

Search moves and search tactics in online searching have been examined by (among others) Bates (1979, 1990); Fidel (1985); Wildemuth (2004), and Xie and Joo (2010). Fidel (1985) described a search move as a query formulation during a search process. By analyzing 90 searches, she identified operational and conceptual search moves. Operational search moves modified the query without changing the meaning of query components, whereas conceptual search moves changed the meaning of query components. Bates (1979) identified 29 search tactics in four categories: monitoring, file structure, search formulation, and term tactics. In subsequent research, Bates (1990) presented levels of search activities from a search move to a search strategy. A search move refers to an identifiable action that is a part of information searching.

According to Wildemuth (2004), a search move refers to an iteration in the search formulation and reformulation process. Instead of analyzing individual search moves, Wildemuth examined the sequences of moves. A set of moves that were temporally and semantically related was described as a search tactic. Xie and Joo (2010) analyzed transitions of search tactics during online searching. They defined search moves as basic actions in the information search process. Search moves form search tactics, which refer to users' choices that are applied to advance the search process.

3.2. Search tactics for fiction

Pejtersen (1989) has identified four search tactics that apply to fiction searching in online catalogs. A bibliographical search tactic is used when readers are searching for a known item or author. An analytical search tactic is used when readers wish to access books about a specific topic, such as the French revolution. A search-by-analogy tactic is generated when readers want something similar to a novel they have read. A browsing tactic is applied in situations when readers have only a vague idea of what they would like to read. To support the four search tactics applied for fiction searching, Pejtersen (1989) presented the BOOKHOUSE model. In the model, the works of literature are indexed with the AMP classification scheme, which has four independent dimensions: subject matter, frame, author's intention, and accessibility (Pejtersen, 1989; Pejtersen & Austin, 1983, 1984). Evaluation of BOOKHOUSE revealed that users were especially satisfied with the indexing principles because they were effectively matching the users' perception of the contents of the books. Also, the ability to use different search tactics gave the users a sense of system flexibility, as it enabled different routes to the contents of the collection. The most popular search tactics used in BOOKHOUSE were the analytical search tactic and search within book cover images (Pejtersen, 1989).

Adkins and Bossaller (2007) compared different entry points for fiction collections in online bookstores, reader advisory databases, and public library catalogs. They suggested that public library catalogs can function effectively, especially for known-item searches. Hinze, McKay, Vanderschantz, Timpany, and Cunningham (2012) observed the book selection process in academic libraries and found that browsing and serendipitous discoveries were not effectively supported in digital libraries.

Oksanen and Vakkari (2012) examined search tactics in an enriched public library catalog. They found that advanced search, exploring result lists, and inspecting book pages were the major search moves for fiction. The findings suggested that effort invested in inspecting search results and book pages instead of querying was an essential factor for finding interesting novels in browsing situations. In their gaze tracking study,

Pöntinen and Vakkari (2013) examined how the metadata of books were examined in an enriched catalog compared to a traditional catalog when searching for fiction. They found that the same metadata elements were inspected as much in both catalogs. The strongest predictor of successful book selection was the duration of visits to the author/title information, suggesting that the content description did not play a crucial role in fiction selection.

In fiction searching, the concept of social browsing (Peters, 2011) refers to the ability to re-orientate browsing by social navigational tools (Tang, Sie, & Ting, 2014). Tang et al. conducted a user study of aNobii, an online book sharing website. It provides three social book-finding tools: browsing books by known authors, browsing similar bookshelves, and browsing friends' bookshelves. The author browsing was found to be the most efficient, while browsing similar and friends' bookshelves produced more serendipitous choices.

In addition to social browsing, user interfaces such as Blended Shelf (Kleiner, Rädle, & Reiterer, 2013) and Bookfish (Pearce & Chang, 2014) visualize book collections and offer the possibility for serendipitous browsing and discovery with visual tools. However, evaluation studies of the usability and usefulness of these tools in fiction book search remain scarce. In the Bohemian Bookshelf, introduced by Thudt et al. (2012), serendipitous book discovery in a library catalog was enabled through information visualization techniques such as virtual book piles, book recommendations, and visualization of connections among books. The Bohemian Bookshelf was created to support especially open-ended explorations and the browsing search tactic. The use of folksonomies, such as tag clouds, and their effect on user behavior in an academic library context was studied by Anfinnsen, Ghinea, and de Cesare (2011). Tag clouds were found to encourage users to browse a catalog in depth and to enhance active user involvement. Similarly, Hassan-Montero and Herrero-Solana (2006) noticed that tag clouds may increase users' browsing activity.

4. Procedures

4.1. Participants

Participants were recruited in three cities situated in the Pirkanmaa, Kanta-Häme, and Uusimaa regions in Finland. In order to find participants with realistic and natural fiction reading interest, 80 people were recruited in public libraries, in fiction reading groups, and in writing and literature classes in the Open University of Finland. The writing and literature classes were open to all adult students paying a course fee, regardless of age or educational background. The literature and writing classes were arranged in cooperation with university departments. In addition, the snowball sampling method and a newspaper advertisement were used. Participants were offered a movie voucher to participate in the study. Participants with no genuine fiction reading interest were excluded from the study.

The participants were randomized into control and test groups, where the control group used a traditional online catalog and the test group used an enhanced catalog that provided enriched fiction related metadata elements (such as cover images and book descriptions) and enabled users to browse the collection via different access points to literature. In both groups, 18% of the participants were male and 82% female. In both groups, the age distribution of participants varied from 20 to 80. In the group using a traditional catalog, participants averaged 34 years of age (SD 12.7). In the case of the enriched catalog, participants averaged 42 years of age (SD 16.8). The age distribution of participants in the enriched catalog group was significantly higher in comparison with the age distribution in the traditional catalog group (t = 2.46, p < 0.05). The educational level of the participants was designated as middle or high. Middle level education is a comprehensive school education, which is finished when the syllabus of a comprehensive school education has been completed. This compulsory education starts in the year of the child's seventh birthday and is finished when the

syllabus of a comprehensive school education has been completed, nine years later. High education refers to an education consisting of at least matriculation examination, which is three to four years of studies after the comprehensive school. In both groups, 18% of the participants had a middle level education and 82% a high level education.

81% of the participants in the traditional catalog group and 58% of the participants in the enriched catalog group used online catalogs for fiction searching at least once every three months. In the traditional catalog group, 5% of the participants had never used online catalogs for fiction searching. In the enriched catalog group, 10% of the participants had never used online catalogs for fiction searching. No significant differences were found in the use of online catalogs between the groups.

4.2. Simulated search tasks

Five simulated search tasks were designed based on previous research (Adkins & Bossaller, 2007; Goodall, 1989; Moyer, 2007; Peters, 2011; Pejtersen, 1989; Ross, 2001; Spiller, 1980; Thudt et al., 2012; Van Riel, 1999; Yu & O'Brien, 1996). In two search tasks, the participants were given an author or a topic with which to begin the search process. The remaining three search tasks explored the concept of individual and dynamic information needs, as the participants were asked to proceed according to their personal preferences without any given topic. Search tasks reflect the following typical search tactics in fiction searching: known author/title search (Mikkonen & Vakkari, 2012; Pejtersen, 1989; Spiller, 1980; Yu & O'Brien, 1996), open ended browsing (Bawden, 2011; Pejtersen, 1989; Thudt et al., 2012), topical search (Bawden, 2011; Pejtersen, 1989; Thudt et al., 2012), search by analogy (Pejtersen, 1989), and searching without conducting a query (Thudt et al., 2012). Simulated search tasks were as follows:

○ The known author search: "A friend of yours recommends you to familiarize yourself with the novels of Olli Jalonen. Find Olli Jalonen's novels and choose two novels which are of interest to you."
○ The topical search: "Find three novels about upper class life in the 19th century."
○ The open-ended browsing: "Find three novels that interest you which you would like to read."
○ The search by analogy: "Think of and mention one novel that you have read and found interesting recently. Now search for three novels that you would consider similarly interesting as the one you mentioned."
○ The open-ended browsing for vacation reading: "You are leaving for a holiday trip. Find three entertaining novels for the trip without using the free text box."

In the task of open-ended browsing for vacation reading, the search began without conducting a query. In the test group, the participants were instructed to begin the search by using the catalog's other features instead of issuing a query. For the control group using the traditional system, the search task was presented in a slightly different way, as it was not possible to browse the collection without first conducting a query. The statement for the fifth task in the control system was: "You are leaving for a holiday trip. Find three entertaining novels for the trip."

4.3. The catalogs used

BookSampo-service (Sampo) was used as the enhanced catalog (the test group) for the study. Compared to the baseline catalog Satakirjastot, Sampo provides fiction related metadata that is more varied and allows for different access points to literature. Sampo contains metadata for the adult fiction collection from the HelMet Web Library, which is the web service of the city libraries of the Helsinki Metropolitan area in Finland. In Sampo, the associations and similarities between the works of literature are realized by semantic web technologies such as the

ontologization of the fiction thesaurus Kaunokki, which is used for fiction indexing in Sampo (Hypén & Mäkelä, 2011).

The front page of Sampo offers a variety of access points to the collection. Users can begin the search process with a basic search (similar to the baseline catalog), by browsing book cover images in a book cover carousel, by skimming other users' virtual bookshelves, or by looking at a subject term cloud (Fig. 1). The book cover carousel on the front page displays 25 randomly picked items which change daily. Further, five randomly picked virtual bookshelves created by other users are displayed on the front page, with an option to view all the user-created bookshelves. Book reviews and book recommendations are provided daily by the administrators of the service, also on the front page.

The search results are sorted in two separate sorts: one by author and another by genre of literature (novels, plays, short stories). A book page represents the content of a particular novel in detail: metadata on subject and themes, milieus, settings, main characters, and time frames are included alongside traditional subject terms and cover images. Book descriptions, automatic recommendations, and text samples are also included.

Satakirjastot (Sata) is the web service of the city libraries of the Satakunta region in Finland. Since it includes a traditional public library online catalog, it was used as the baseline system. The service consists of a library catalog, information retrieval system, customer record, basic and advanced search from a variety of databases, and links to each library's own web services.

In Sata, the metadata for fiction includes bibliographic information and subject terms from the fiction thesaurus Kaunokki (Saarti & Hypén, 2010). For books published recently, cover images and book descriptions are also available. The search options include basic and advanced search. Basic search matches the given words in every accessible part of the data. In advanced search, it is possible to narrow the search by several options, including specific metadata fields, in the dropdown menu (Fig. 2). The search results are sorted by format (such as book, DVD, CD). After narrowing down to the preferred format, the search results can be sorted by author, title, publication year, or class.

When the catalogs for the experiment were considered, the aim was to identify a traditional catalog with a fiction collection comparable to Sampo in size and content. The collection in Sata was found to be relatively similar to Sampo's collection. The fiction collection of Sata consisted of 77,000 titles and the fiction collection of Sampo of 73,000 titles, both including novels, plays, and poems. Both collections included the classics and recently published titles of Finnish and foreign literature. Before the experiment, the researchers conducted several test queries in both catalogs in order to compare the results. They were found to share several overlapping titles. The analysis of Sata and Sampo revealed that the size and content of these collections were comparable, and therefore they could be used in the experiment.

4.4. Experimental procedure

Before conducting the user tests, the experimental setting was pretested with one participant to gain information on the duration of the test as a whole and to see whether the instructions were unambiguous enough. The experiment consisted of the following steps:

1. Pre-questionnaire including demographic questions, questions concerning participants' search experience in online catalogs in general, participants' familiarity with the catalog used in the experiment, and their reading interest.
2. Introduction to the experiment.
3. Brief demo of the retrieval system (approximately two minutes long).
4. Execution of five search tasks.
5. A post-task questionnaire after each completed search task.

Fig. 1. BookSampo book cover carousel (1) and subject term cloud (2).

6. A post-session questionnaire after the completion of all search tasks.
7. Brief post-task interview after one search task (decided in advance).

The time for completing the tasks was not limited. As mentioned above, the participants were randomized into control and test groups. Each participant completed the tasks either on Sampo or Sata. The pre-questionnaire confirmed that the participants in both groups were unfamiliar with the catalogs used in the experiment. The known-author task functioned as a training task and was conducted first with each participant. Latin square rotation was used with the four other tasks.

Each participant completed the tasks individually. During the experiment, the researcher was present to help in case technical problems occurred. However, the researcher did not guide or help the participants in questions concerning the completion of the search tasks.

In each task, the participants were asked to search for three novels that were of interest to them except in the known author task, in which the participants were asked to find only two interesting novels. The known author task functioned as a training task at the beginning of the experiment. It was thought that finding two interesting novels was enough to familiarize the participants with the catalog's basic functions. After each search task, the participants were asked to rank the novels found according to how much they were of interest to them using an ordinal scale ranging from 1 to 3, where 1 was "a little interesting," 2 was "somewhat interesting," and 3 was "very interesting." A score of 0 was used if an interesting novel was not found.

4.5. Analysis

4.5.1. Measures

The search logs were saved with Morae-software. The search moves applied in fiction searching in five search tasks were analyzed (Table 1). A search move refers to an identifiable basic action during the search process (similar to work done by Bates, 1990; Xie & Joo, 2010). Search moves were identified by examining the recorded user tests. In Sampo, the label pivot browsing (Peters, 2011) was applied to search moves that re-orientated browsing to follow features such as virtual bookshelves, a tag cloud, a cover image carousel, and automatic recommendations. In Sata, pivot browsing was not possible.

After identifying the individual search moves, a coding scheme was developed to measure the search moves. The scheme was slightly different between the two catalogs used as they offer different search features. The identified search moves were measured by their frequency and duration. Success in search tasks was measured by the ratings of the interesting novels found given by participants. The average book scores per task were calculated. Each of the search tasks was coded individually. In addition, the measures of the search session were also coded. The coding scheme is presented in Table 2.

4.5.2. Statistical analysis

Variables presented in the coding scheme were screened for normality. The variables were not normally distributed (Shapiro–Wilk = .95,
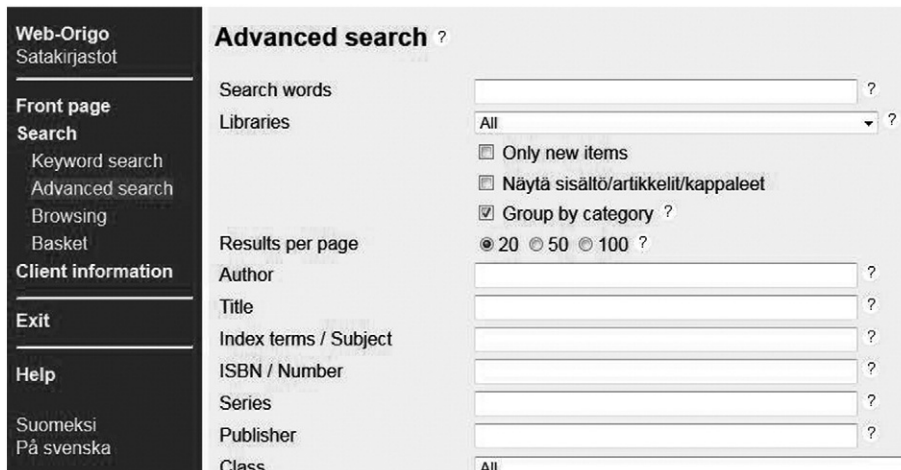

Fig. 2. Advanced search in Satakirjastot.

Table 1
Search moves in both catalogs.

Search moves in Sata	Search moves in Sampo
Conducting queries in basic searching	Conducting queries in basic searching
Conducting queries in advanced searching	Conducting queries within book cover images
Conducting queries in browsing searching	Examining the search result list
Examining the search result list	Exploring an individual book page
Exploring an individual book page	Exploring the catalog's front page (containing subject term cloud, book cover carousel, book recommendations)
Reading description of a novel (in a separate window)	
Examining book cover of a novel (in a separate window)	Pivot browsing (examining automatic recommendations in a book page, browsing book cover images, browsing other users' virtual book piles, reading book recommendations produced by the maintenance

p < 0.05 in each variable). Shapiro–Wilkinson's test was used, as there were fewer than 50 cases (Hair, Anderson, Tatham, & Black, 1998). Friedman's test was used to test differences in the use of search moves between search tasks within a catalog. The known author task differed from the other search tasks, as the participants were asked to find two interesting novels compared to three novels in other search tasks. To enable the comparison of search moves between tasks within a catalog, the frequency and duration variables of the known author task were multiplied with the coefficient of 1.5 in basic statistics and Friedman's tests.

When comparing the search moves between catalogs, the variables were again screened for normality. Variables were not normally distributed (Kolmogorov–Smirnov = .95, p < 0.05 in each variable). Kolmogorov–Smirnov's test was used, as there were over 50 cases (Hair et al., 1998). The Mann–Whitney nonparametric U-test was used to test significant differences in the use of search moves between catalogs.

5. Findings

Significant differences in task completion times between the catalogs were not found (Table 3). In both catalogs, the searchers devoted most time to select novels in browsing tasks where they proceeded according to their genuine reading interest. In Sata, the known author task was completed most rapidly. Significant differences were found between the known author task and the other tasks (Friedman's between tasks test-values Known-Topical 6.4, p < .05, Known-Open 28.9, p < .001, Known-Analogy 16.9, p < .001, Known-Vacation 32.4, p < .001). In Sampo, the topical search was completed the most rapidly. The differences between the topical search and the other tasks were significant (Friedman's between tasks test-values Topical-Open 16.0, p < .001, Topical-Analogy 6.4, p < .05, Topical-Vacation 5.8, p < .05).

Table 2
The coding scheme.

Duration-variables	Time to complete a search task (Sata, Sampo)
	Time to submit queries (Sata, Sampo)
	Time to formulate the first query (Sata, Sampo)
	Time to formulate the following queries (Sata, Sampo)
	Time to view the search results (Sata, Sampo)
	Time to view the book pages (Sata, Sampo)
	Time to view the cover images (Sata, Sampo)
	Time to read the description of the book (Sata)
	Time to view the catalog's front page (Sampo)
	Time to pivot browsing actions (Sampo)
Frequency-variables	The number of opened book pages
	The number of queries
	The number of individual search moves

Table 3
Task completion times. (min.sec).

Search type	Sata (n = 40) M (SD)	Sampo (n = 40) M (SD)
Search session	39.55 (12.37)	44.27 (13.04)
Known author	5.37 (2.39)	6.30 (3.14)
Topical search	6.35 (3.09)	5.11 (2.56)
Open ended	7.45 (3.17)	8.22 (3.52)
Search by analogy	7.32 (4.43)	8.20 (4.47)
Vacation reading	7.50 (3.32)	7.06 (3.51)

Across all tasks, more effort in the form of queries and search moves was needed in Sata compared to Sampo for searchers to select interesting novels (Table 4). The number of queries was substantially greater in Sata compared to Sampo, especially in browsing tasks. In Sampo, some of the participants proceeded without a query in the browsing tasks, while in Sata, it was necessary to issue a query in order to do any browsing.

The number of search moves was significantly higher in Sampo compared to Sata. In both catalogs, the known author task was completed with fewer search moves compared to other tasks. Most search moves were needed to select novels in open-ended browsing and search by analogy in both catalogs. There were no significant differences in the number of opened book pages between catalogs. This hints that in Sata, more effort was needed in form of queries and search moves to view an equivalent number of title pages to Sampo.

In Sata, differences between search tasks were found in the number of queries and opened book pages. Significantly fewer queries were made in the known author task compared to other search tasks (Friedman's between tasks test-values Known-Topical 27.13, p < .001, Known-Open 26.47, p < .001, Known-Analogy 19.0, p < .001, Known-No query 29.43, p < .001). Although the differences were nonsignificant, more book pages were opened in the known-author task compared to other tasks. In the known-author task, some of the participants selected the novels after glancing at each item in the results list. In other search tasks, solely the items that the participants were intrigued by were opened for examination. This may partly explain the difference in the number of opened title pages between tasks.

In Sampo, the greatest difference between search tasks occurred in the number of queries. The task open-ended browsing for vacation reading was completed with significantly fewer queries compared to other tasks (Friedman's between tasks test-values No query-Known 17.0, p < .001, No query-Topical 22.5, p < .001, No query-Open 19.59, p < .001, No query-Analogy 28.44, p < .001). In the open-ended browsing for vacation reading, the participants were asked to use Sampo's other features instead of querying. Thus, the majority of the participants selected novels for vacation without a single query. Also in the open-ended browsing, 28% of the participants proceeded without a single query. Search tasks in Sampo differed also in the number of search moves. The known-author task required the least moves while search by analogy required the most. This difference was significant (Friedman's test-value 16.89, p < .001).

To further examine the differences in search moves between search tasks and catalogs, the time used in the most popular search moves in both catalogs was analyzed. In Sata, options for submitting queries were basic search, browsing, and advanced search. In Sampo, solely the basic search option was provided. In order to compare the time devoted to submitting queries between catalogs, a query variable was calculated. In Sata, the query variable consists of the time used for basic, browsing, and advanced searching. In Sampo, the query variable consists of the time used for basic searching. In Sampo, multiple entry points to the collection and possibilities to pivot browsing were provided at the interface's front page. A variable named as "front page" was calculated to measure the time searchers devoted to examining Sampo's front page and the time users invested in pivot browsing actions. In Sata, pivot browsing was not possible and the only option for browsing the

Table 4
Basic statistics of queries, book pages, and moves.

Search type	Sata (n = 40) M			Sampo (n = 40) M			Mann–Whitney U-test between catalogs p		
	Queries	Book pages	Moves	Queries	Book pages	Moves	Queries	Book pages	Moves
All	18	26	112	10	25	81	**.000**	.996	**.001**
Known	2	7	22	2	5	14	.956	.300	**.029**
Topical	4	5	24	3	5	16	**.030**	.121	**.001**
Open	4	6	26	2	5	19	**.001**	.801	**.048**
Analogy	4	5	26	3	7	21	.054	.087	.218
Vacation	5	4	24	1	5	16	**.000**	.412	**.001**

Bold values indicate significance at <.05.

collection was to issue a query. Hence, the front page variable was calculated solely for Sampo.

Significant differences were found in time devoted to issuing queries, examining search results, and inspecting book pages between catalogs (Table 5). Across all tasks, searchers formulated queries significantly longer in Sata compared to Sampo. This hints that in Sata, when selecting interesting novels, it is productive to invest time in choosing keywords and consequently gain interesting search results for browsing. In Sampo, a single box search option might support quick querying. In addition, other access points to the collection may reduce the need for querying, which may explain why the time invested in querying was shorter in Sampo. In Sata, more time was invested in examining search results, but less time in viewing book pages compared to Sampo.

In Sata, from the three search options provided, users preferred the advanced search option. The findings show that queries were issued significantly more quickly in the known-author task compared to the other tasks (Friedman's between tasks test-values Known-Topical 36.1, p < .001, Known-Open 28.9, p < .001, Known-Analogy 22.5, p < .001, Known-No query 36.1, p < .001). Across all tasks, more time was devoted to search results compared to book pages in Sata. The most time was devoted to search results in topical browsing (Friedman's between tasks test-values Topical-Open 6.4, p < .05, Topical-Analogy 4.9, p < .05).

In Sampo, the greatest difference in search moves between tasks occurred in time devoted to querying. As with Sata, queries were issued significantly more quickly in the known author task compared to other tasks (Friedman's between tasks test-values Known-Topical 13.56, p < .001, Known-Analogy 11.31, p < .05, Known-No query 25.6, p < .001). Also in the vacation-reading task, query time was limited. This is due to the task instructions that asked participants to use Sampo's other entry points instead of issuing a query. Nevertheless, a few participants issued queries regardless of the instructions in the vacation reading task. Differences occurred also in time devoted to search results and book pages. Users devoted the most time to search results in the search by analogy (Friedman's between tasks test-values Known-Analogy 8.1, p < .05, Topical-Analogy 3.6, p > .05, Open-Analogy 10.0, p < .05, Vacation-Analogy 16.9, p < .001). Book pages were examined the most in the open-ended browsing task (Friedman's between tasks test-values Known-Open 9.26, p < .05, Topical-Open 2.5, p > .05, Analogy-Open 4.33, p < .05, Vacation-Open 0.03, p > .05).

In Sampo, time to examine the catalog's front page and to conduct pivot browsing actions seemed to be associated to the task type. In tasks where the author or the topic was pre-decided, searchers tended to prefer a similar search approach as in Sata: querying and result list examination over viewing book pages. Growth in the exploratory and serendipitous nature of the search situation seemed to increase the time users devoted to Sampo's various entry points to the collection instead of querying. In the open-ended browsing and the vacation reading, users devoted the most time to the catalog's visual and social navigational tools. Differences were very significant (Friedman's between tasks test-values Open-Known 20.57, p < .001, Open-Topical 17.07, p < 001, Open-Analogy 16.13, p < .01, Vacation-Known 26.0, p < .01, Vacation-Topical 27.5, p < .001, Vacation-Analogy 18.8, p < .001). This hints that in Sampo, when browsing for good novels to read, it is productive to invest time in viewing the enriched front page and following the social and visual features for clues as to possible interesting titles.

In general, participants were successful in finding interesting novels in both catalogs. Each participant selected at least one interesting novel across all search tasks. The majority of the participants selected three novels in each search task in both catalogs. Significant differences in average book scores between catalogs were not found (Table 6). In both catalogs, the known-author task gave the poorest results. In both catalogs, the differences between the known-author task and the browsing tasks were very significant (Friedman's test-values Sampo Known-Open 40.0, p < .001, Known-Analogy 36.1, p < .001, Known-Vacation 31.4, p < .001, Sata Known-Open 40.0, p < .001, Known-Analogy 40.0, p < .001, Known-Vacation 36.1, p < .001), and significant between the known author task and the topical search (Friedman's test-value Sampo 5.4, p < .05, Sata 6.4, p < .05). In both catalogs, the highest scores were given to the novels found in the three browsing tasks, where the participants proceeded according to their genuine interest. It seems that in both catalogs, the participants were the least satisfied with the novels found in tasks where author or topic was pre-decided by the researcher.

6. Discussion

6.1. Differences in search moves between the catalogs

More queries, search moves, and opened book pages were required in Sata to reach equivalent search success as in Sampo. Also, the results

Table 5
Time used in the most popular search moves (min.sec).

Search type	Sata (n = 40) M			Sampo (n = 40) M				Mann–Whitney U-test, between catalogs p		
	Query	Search results	Book pages	Query	Search results	Book pages	Front page	Query	Search results	Book pages
All	10.18	15.57	6.35	3.6	10.2	9.24	2.15	**.000**	**.000**	**.012**
Known	0.28	3.37	1.32	0.36	2.57	2.18	0.08	**.021**	.194	.378
Topical	2.20	2.49	1.26	0.55	2.11	1.22	0.10	**.000**	.183	.836
Open	2.22	3.34	1.29	0.42	2.00	3.10	1.55	**.000**	**.000**	**.027**
Analogy	2.24	3.40	1.28	0.52	3.16	2.35	0.24	**.000**	.222	**.008**
Vacation	3.03	3.36	1.11	0.12	1.21	1.00	2.39	**.000**	**.000**	**.013**

Bold is used in the Table when p-value is statistically significant.

Table 6
Average book scores per task.

Search type	Sata (n = 40) M (SD)	Sampo (n = 40) M (SD)
Known author	3.5 (0.7)	3.7 (0.7)
Topical	4.7 (1.8)	4.4 (1.6)
Open	6.0 (0.8)	6.0 (0.7)
Analogy	6.1 (0.7)	5.0 (0.8)
Vacation	6.0 (1.0)	5.7 (1.0)

demonstrated that in Sata, a typical search strategy for interesting titles seemed to be issuing queries and considering suitable entry terms carefully and devoting more attention to search results instead of book pages. Finally, in Sampo, time devoted to exploring the catalog's enriched front page and multiple entry points together with attention to the enriched result list seemed to be a common search approach.

6.1.1. Access points to the collection

In a traditional catalog that does not support browsing with visual or social navigational tools, a query is a necessity for browsing the collection. An essential precondition for a satisfying query is a knowledgeable literate user who is capable of typing suitable author or title names or desired themes as keywords. The capacity to retrieve an interesting result list was a crucial condition for successful discoveries in Sata, compared to browsing cover images, virtual book piles, and the enriched front page in Sampo. Hence, more emphasis was devoted to querying in Sata compared to Sampo.

While the single option for accessing the collection in Sata was querying, Sampo provided multiple entry points immediately at the front page. In Sampo, exploring the tag cloud, virtual book shelves, and the book cover carousel at the enriched front page offered an appealing starting point for the search, perhaps by causing curiosity, leading users to explore the catalog more closely. This might explain why time devoted to the catalog's front page seemed to be a typical search approach in Sampo. This is in line with the finding by Anfinnsen et al. (2011) that tag clouds encouraged users to browse a catalog in depth and enhanced active user involvement. Buchanan and McKay (2011) noticed that when readers were expressing their reading wishes in query terms, lack of knowledge was sometimes an obstacle to successful discoveries. Thus, the social and visual navigational tools that enable browsing without querying might have increased the possibility of serendipitous discoveries in Sampo. As Hassan-Montero and Herrero-Solana (2006) point out, querying requires the user to formulate information needs, whereas visual browsing allows the user to recognize information needs while scanning the interface. Searching without querying turned out to be a popular search approach, especially in browsing tasks in Sampo. When queries were issued in Sampo, they were formulated significantly more quickly than in Sata. This finding might support the idea that Sampo's use of an ontologized fiction thesaurus better supports quick and simple querying, as this method enriches the results produced for simple queries by exploiting connections among items.

6.1.2. User behavior on search result pages

In Sata, the identification of an interesting title from a search result page might have been challenging for users, as only author and title names were provided. This was even harder when the author was previously unknown. In Sampo, the result list was enriched, containing a cover image and a snippet of a blurb together with author and title names. Across all tasks, the users of Sata devoted more time to search results compared to Sampo. The enriched result list in Sampo might have helped searchers identify possibly interesting titles more efficiently and more quickly than in Sata. With the rich information on the result list, the users of Sampo might have made pre-selections more quickly from the result pages, which may explain why the time invested in search results was shorter in Sampo compared to Sata. This agrees with Vakkari and Pöntinen (2015), who found that an enriched catalog

helped users identify potentially clickable items on the result list sooner as compared to a traditional catalog.

6.1.3. User behavior on book pages

Similar to work by Pöntinen and Vakkari (2013), the present study found that more time was devoted to book pages across all tasks in Sampo compared to Sata. In Sata, information on novels' topics and themes was discovered on book pages by inspecting subject terms and publication information. A description and a cover image were provided with novels published recently, though in a separate window. Occasionally, the information on book pages was fairly limited, as subject terms were not provided. In Sampo, the book pages were always enhanced with many subject terms, a detailed description of the plot, a cover image, and automatic recommendations of similar titles. The book page display in Sampo may have intrigued the users more than in Sata, which may possibly explain the difference in time devoted to book pages between catalogs. Users of Sata opened more book pages to select equally interesting titles as compared to users of Sampo. This might partly be a result of Sampo's enriched result list, as opening book pages was not necessary to identify appealing titles.

6.2. Differences in search moves between tasks

The most popular search moves across all search tasks in both catalogs were querying, search results exploration, and book page inspection. This resembles the results in Oksanen and Vakkari (2012).

In Sata, the major difference between search tasks occurred in time invested in querying. Queries were conducted significantly more quickly in the known-author task compared to other tasks. This suggests that participants were used to searching for novels with author names, while querying with themes or topics might have been more unfamiliar to users, requiring more time. Interestingly, users devoted less time to search results in the topical search compared to other search tasks. This seems to reflect the limited information provided on search result pages: a novel's topicality could not be detected from search results list if the title was previously unknown.

In Sampo, queries were issued quickly across all tasks. As with Sata, the least time was invested in querying for the known-author task (the task open-ended browsing for vacation reading excluded, as the participants were asked not to issue queries during this task). Interestingly, users devoted the most time to search result pages during the known-author and search-by-analogy tasks. It seems that the enriched result list helped users to identify interesting titles, especially when the author was pre-decided. Also, users seemed to assess similarity between novels by inspecting the enriched results list, possibly by looking at the snippet of a blurb. Time devoted to the catalog's front page and time to pivot browsing seemed to be associated with the task type: users preferred pivot browsing and searched for clues of interesting titles at Sampo's front page more when the search situation was explorative by nature. When the author or the topic was pre-decided, the front page was given less attention.

The highest average book scores in both catalogs were produced by the browsing tasks where the participants proceeded according to their genuine interest. This hints that in both catalogs, the participants were the most satisfied with novels found when authors, titles, and topics were decided by the participants themselves.

To conclude, even though differences occurred in users' search behaviors between catalogs, in general, the users of both catalogs were successful in selecting interesting titles. In a traditional catalog a) a query is a necessary condition for browsing results, and b) returning to the search results from a book page is a necessary condition for proceeding to the next interesting item. Thus, in a traditional catalog, in both known-author and browsing tasks, effort needs be invested especially in well-designed queries, examining search results for clues of interesting topics, and exploring book pages in detail. A traditional catalog performs well for tasks with a known author, title (see Adkins &

Bossaller, 2007), or topic (with advanced search). In an enriched catalog, the visual front page and the automatic recommendations enable the participants a) to browse without querying, and b) to proceed from item to item without the need to return to the search results. An enriched catalog performs well for open-ended browsing tasks that aim for serendipitous discovery.

6.3. System design

A few suggestions on system design can be made based on these findings. An appealing starting page in Sampo turned out to be a major distinguishing factor between catalogs. Also, time devoted to exploring visual and social navigational tools on Sampo's front page seemed to be a favored search approach, especially in explorative search situations. Thus, providing an appealing starting point for the search that piques curiosity and encourages users to explore the catalog thoroughly would enhance traditional online catalogs.

The results suggest that issuing queries other than for a known author or title was somewhat problematic for users. When users rely on their pre-knowledge, the books selected tend to be from well-known authors and genres. Providing multiple access points for the collection would support selection of novels with desired themes or topics and enhance new and unusual book selections.

Querying is a fruitful and common search approach, especially in tasks where the author or title is already known. In Sampo, the missing advanced search option would have been useful for some tasks. However, in Sampo, simple queries produced rich results, as the ontology allowed for query expansion. Thus, the optimal interface would provide both a single search box and an advanced search option together with an ontologization of a thesaurus. When searching for a certain topic, the participants tended to associate query terms (query terms were manually collected from recordings) with different dimensions of the novels, e.g., theme and emotion. The indexing of novels with both objective and subjective index terms (Adkins & Bossaller, 2007) would support topical and thematic search.

In Sampo, automatic recommendations were used when searching for novels similar to previously read ones. However, similar to findings by in Pöntinen and Vakkari (2013), the recommendations often did not catch participants' interest as they were positioned in the right-hand corner of the book page. In Sata, fuller book descriptions and book cover images were positioned in the right corner of the book page but were still infrequently used during the experiment. Thus, when developing systems for fiction retrieval, both the positioning and the visual appearance of the recommendations, descriptions, and cover images should be designed to catch the eye. Book descriptions should be available on the book page, not behind a click on a separate window.

Even though the results gave some hints of the association between search behavior and search success, there continues to be much more room for further studies identifying factors associated with successful discovery in online catalogs. For example, the associations of factors for searching such as reading preferences, reading diversity, and emotions remain unknown. The use of search moves in this study revealed interesting differences in fiction readers' search behaviors between two types of catalogs. In future research, search moves will be further examined during query reformulation intervals and used to elicit information about readers' document viewing behavior. The association of search moves and search success will also be further analyzed between successful and unsuccessful query reformulation intervals. Search moves will also be compared between different types of fiction readers to further understanding about book selection behavior among readers with varied literary preferences.

6.4. Limitations

The sample was biased toward females and the highly educated. Also, the participants conducted simulated search tasks in an experimental setting. Thus, it is not reasonable to generalize the findings to the population at large. Instead, the transferability of findings could be highlighted. As the simulated search tasks reflected the major search tactics for fiction retrieval, it could be assumed that the search moves of fiction readers in another setting would share some characteristics compared to these results.

7. Conclusion

Understanding of fiction readers' search behavior in online catalogs will be essential as fiction collections become wholly digital and continue to grow. This study explores the basic characteristics of fiction book searching in various search scenarios in online catalogs. The study revealed that fiction readers' search behaviors in a traditional catalog and an enriched catalog differ but may result in equal success. Providing multiple appealing entry points for fiction collections together with an enriched results list may result in successful, serendipitous discoveries. Furthermore, providing visual and social navigational tools may trigger ideas for search and consequently support the discovery of interesting new material. Designing user-friendly interfaces for fiction readers with varied literary interests and searching competencies requires knowledge from large-scale evaluative user studies. The results of this study offer a solid starting point for further examination of user behavior in online catalog fiction discovery and are a significant step toward a comprehensive understanding of search behaviors for casual-leisure reading material.

References

Adkins, D., & Bossaller, J.E. (2007). Fiction access points across computer-mediated book information sources: A comparison of online bookstores, reader advisory databases, and public library catalogs. *Library and Information Science Research, 29*, 354–368.

Anfinnsen, S., Ghinea, G., & de Cesare, S. (2011). Web 2.0 and folksonomies in a library context. *International Journal of Information Management, 31*, 63–70.

Bates, M.J. (1979). Information search tactics. *Journal of the American Society for Information Science, 30*, 205–214.

Bates, M.J. (1990). Where should the person stop and the information search interface start? *Information Processing & Management, 26*, 575–591.

Bawden, D. (2011). Encountering on the road to Serendip? Browsing in new information environments. In A. Foster, & P. Rafferty (Eds.), *Innovations in information retrieval: perspectives for theory and practice* (pp. 1–22). London, UK: Facet Publishing.

Buchanan, G., & McKay, D. (2011). In the bookshop: Examining popular search strategies. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, 11*, 269–278.

Eden, B. (2002). Using metadata to build an enriched library catalog. *Library Technology Reports, 38*, 65–70.

Fidel, R. (1985). Moves in online searching. *Online Review, 9*, 61–74.

Goodall, D. (1989). Browsing in the public libraries. *(LISU Occasional Paper, 1)*Loughborough, UK: Library and Information Statistics Unit.

Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1998). *Multivariate data analysis.* Upper Saddle River, NJ: Prentice-Hall.

Hassan-Montero, Y., & Herrero-Solana, V. (2006). Improving tag-clouds as visual information retrieval interfaces. *Proceedings of the 1 International Conference on Multidisciplinary Information Sciences and Technologies, October 25–28, 2006, Merida, Spain.* Retrieved from http://www.yusef.es/improving_tagclouds.pdf.

Hessel, H., & Fransen, J. (2012). Resource discovery: Comparative survey results on two catalog interfaces. *Information Technology and Libraries, 31*, 21–44.

Hinze, A., McKay, D., Vanderschantz, N., Timpany, C., & Cunningham, S.J. (2012). Book selection behavior in physical library: Implications for ebook collections. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, 12*, 305–314.

Hypén, K., & Mäkelä, E. (2011). An ideal model for an information system for fiction and its application: Kirjasampo and semantic web. *Library Review, 60*, 279–292.

Kleiner, E., Rädle, R., & Reiterer, H. (2013). Blended shelf: Reality-based presentation and exploration of library collections. *Extended Abstracts: The ACM Conference on Human Factors in Computing Systems, 31*, 577–582.

Lau, E.P., & Goh, D.H.L. (2006). In search of query patterns: A case study of a university OPAC. *Information Processing and Management, 42*, 1316–1329.

Maker, R. (2008). Finding what you're looking for: A reader-centered approach to the classification of adult fiction in public libraries. *Australian Library Journal, 57*, 169–177.

Mikkonen, A., & Vakkari, P. (2012). Readers' search strategies for accessing books in public libraries. *Information interaction in context symposium, 4*. (pp. 214–223).

Moulaison, H.L. (2008). OPAC queries at a medium-sized academic library: A transaction log analysis. *Library Resources and Technical Services, 52*, 230–237.

Moyer, J. (2007). Learning from leisure reading: A study of adult public library patrons. *Reference & User Services Quarterly, 46*, 66–79.

Moyer, J., & Weech, T.L. (2005). The education of public librarians to serve leisure readers in the United States, Canada and Europe. *New Library World, 106*, 67–79.

National Endowment for the Arts (2012). *How a nation engages with art: Highlights from the 2012 survey of public participation in the arts.* National Endowment for the Arts Research Report no. *57,* Washington, DC: National Endowment for the Arts Retrieved from http://arts.gov/publications/highlights-from-2012-sppa.

Oksanen, S., & Vakkari, P. (2012). Emphasis on examining results in fiction searches contributes to finding good novels. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, 12,* 199–202.

J. Pearce, J., & S. Chang, S. (2014). *Exploration without keywords: The Bookfish caseProceedings of the Australian Computer-Human Interaction Conference, 26,* 176–179.

A.M. Pejtersen, A.M. (1989). *The BOOKHOUSE: Modelling user's needs and search strategies as a basis for system design. (Risø Report M-2794)*Roskilde, Denmark: Roskilde Risø National Laboratory.

Pejtersen, A.M., & Austin, J. (1983). Fiction retrieval: Experimental design and evaluation of a search system based on users' value criteria (part 1). *Journal of Documentation, 39,* 230–246.

Pejtersen, A.M., & Austin, J. (1984). Fiction retrieval: Experimental design and evaluation of a search system based on users' value criteria (part 2). *Journal of Documentation, 40,* 25–35.

Peters, I. (2011). Folksonomies, social tagging and information retrieval. In A. Foster, & P. Rafferty (Eds.), *Innovations in information retrieval: Perspectives for theory and practice* (pp. 85–116). London, UK: Facet Publishing.

Pöntinen, J., & Vakkari, P. (2013). Selecting fiction in library catalogs: A gaze tracking study. *Proceedings of the International Conference on Theory and Practice of Digital Libraries, 17,* 72–83.

C.S. Ross, C.S. (2001). *Making choices: What readers say about choosing books to read for pleasure* (pp. 5–21)*The Acquisition Librarian, 13*(25), 5–21.

Saarti, J., & Hypén, K. (2010). From thesaurus to ontology: The development of the Kaunokki Finnish fiction thesaurus. *The Indexer, 28,* 5–21.

Spiller, D. (1980). The provision of fiction for public libraries. *Journal of Librarianship and Information Science, 12,* 238–266.

Stebbins, R.A. (2007). *Serious leisure.* New Brunswick, NJ: Transaction Publishers.

Tang, M.C., Sie, Y.J., & Ting, P.H. (2014). Evaluating books finding tools on social media: A case study of aNobii. *Information Processing & Management, 50,* 54–68.

The Reading Agency (2013). Library facts. Retrieved from http://readingagency.org.uk/news/library-facts004/

Thudt, A., Hinrichs, U., & Carpendale, S. (2012). The Bohemian bookshelf: Supporting serendipitous book discoveries through information visualization. In J. A, Konstan (Chair), Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, May 5–10, 2012 (pp. 1461–1470). New York, NY: ACM.

Vakkari, P., & Pöntinen, J. (2015). Result list actions in fiction search: An eye-tracking study. *Proceedings of the IEEE/ACM Joint Conference on Digital Libraries, 15,* 7–16.

R. Van Riel, R. (1999). *A book to match your mood: Forager information system helps readers select books based on psychological attributes* (pp. 292–293)*Library Association Record, 101,* 292–293.

Van Riel, R., Fowler, O., & Downes, A. (2008). *The reader-friendly library service.* Newcastle upon Tyne, United Kingdom: Society of Chief Librarians.

Waller, V. (2010). Accessing the collection of a large public library: An analysis of OPAC use. *Library and Information Science Research Electronic Journal, 20,* 1–27.

Wildemuth, B.M. (2004). The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science & Technology, 55,* 246–258.

Xie, I., & Joo, S. (2010). Transitions in search tactics during the web-based search process. *Journal of the American Society for Information Science & Technology, 61,* 2188–2205.

Yu, L., & O'Brien, A. (1996). Domain of adult librarianship. *Advances in Librarianship, 20,* 151–190.

Anna Mikkonen is a PhD student at the School of Information Sciences, University of Tampere, Finland. Her research interests include user behavior in fiction book search in digital libraries, and she has published in Proceedings of Information Interaction in Context 2012 and 2014 and in Proceedings of Joint Conference in Digital Libraries 2015.

Pertti Vakkari is a professor emeritus of information studies at the School of Information Sciences, University of Tampere, Finland. His research interests include fiction searching, information seeking, the growth of theories in information studies, subject knowledge, and query expansion based on relevance feedback, task-based information searching, and use of digital libraries. His work has appeared in the proceedings of numerous conferences, in edited monographs, and in journals such as Information Processing & Management, Information Research, Journal of Documentation, and Journal of the American Society for Information Science and Technology.

Books' Interest Grading and Fiction Readers' Search Actions During Query Reformulation Intervals

Anna Mikkonen
University of Tampere
School of Information Sciences
Tel + 358-405301815
anna.mikkonen@uta.fi

Pertti Vakkari
University of Tampere
School of Information Sciences
Tel +358-505288354
pertti.vakkari@uta.fi

ABSTRACT

We compared fiction readers' search actions during various query reformulation intervals. We aimed to understand how readers' search actions differed between successful and unsuccessful QRIs and which search actions predicted the selecting of very interesting novels compared to less interesting ones. We conducted a controlled user study with 80 participants searching for interesting novels. Three types of browsing tasks and two types of catalogs were used. Our results demonstrated that browsing task type was associated to readers' document viewing behavior in terms of observed search result pages, opened book pages and dwell time on book pages. When browsing for topical novels, most effort was required to select somewhat interesting novels. When browsing for good novels, most effort was required to select very interesting ones. Logistic regression analysis yielded that the most significant predictors of higher document value were the number of observed search result pages and opened book pages.

Categories and Subject Descriptors

H.3.7 [**Information Storage and Retrieval**]: Digital Libraries – *systems issues, user issues.*

General Terms

Performance, Experimentation, Human Factors.

Keywords

Fiction retrieval; search actions; query reformulation intervals; online catalogs.

1. INTRODUCTION

In recent years research interest towards interactive information retrieval focusing on causal-leisure situations has increased [e.g. 6, 34]. Searching of fiction is an essential part of casual-leisure search: the search process is not triggered by an explicit information need to be solved [34]. Instead, it begins often with a serendipitous need to find *"just good books to read"*.

Fiction is typically accessed in public libraries [28, 29]. Fiction readers make up a large portion of the public library clientele. In addition to physical library collections, fiction is increasingly searched for and accessed in public libraries' e-collections [29]. Public libraries are actively focusing on increasing the availability

of e-books both in the U.S. and in Europe. [i.e. 27, 29]. Given the growing availability of fiction books in public libraries' e-collections, a good understanding of fiction readers' search behavior in digital libraries is essential to create functioning interfaces for fiction.

In recent years public libraries' online catalogs have been actively developed and enriched to meet the needs of diverse reader groups. Previous studies [i.e. 4, 17, 20] have revealed some major search tactics applied to fiction search in public libraries. However, we have only a few studies focusing on browsing for fiction in online library catalogs. Pöntinen & Vakkari [22] have examined how readers explore the fiction metadata in a gaze tracking study. Jiang [8] has studied social library system users' information seeking modes. Tang et al. [25] have examined book-finding tools on social media.

The purpose of this study is to examine fiction readers' search behavior in browsing tasks during query reformulation intervals (QRIs). QRI refers to users' search actions that take place between two queries [11]. According to Liu et al. [11] issuing a query represents a searcher's cognitive decision to modify the current search tactic, and is likely to affect the searcher's retrieval actions. The search actions between two queries form units, which can be used e.g. for eliciting information about users' document viewing behavior. The aim is to understand how fiction readers' search actions differ between successful and unsuccessful QRIs, which search actions are emphasized when interesting novels are selected, and which search actions predict the finding of very interesting novels compared to less interesting novels.

To address these objectives, we conducted an experiment with fiction readers working on different browsing tasks to find novels. The study addresses the following research questions:

RQ 1. Do the search actions during successful and unsuccessful query reformulation interval types differ a) between an enriched catalog compared to a traditional catalog? b) between browsing tasks?

RQ 2. Are the search actions during query reformulation intervals associated to books' interest grading in browsing tasks? If yes, how?

The study provides information on fiction readers' search behavior in casual-leisure browsing situations. In addition, the study enhances understanding of the factors which are associated to search success when searching for novels.

2. RELATED RESEARCH

Document viewing and usefulness. *Search action* refers to an identifiable basic action during the search process [2, 33]. Liu et al. [11] examined users' document viewing behavior, and predicted document usefulness during QRI. Similarly to [11], in this study QRI was defined as *"an interval from the point when a*

user submits a query to the point when a user starts to enter a subsequent query during the same search session" [11].

[11] noticed that QRIs with useful pages were significantly longer compared to QRIs without useful pages. Moreover, users visited more content and search result pages (SERPs) and spent more time on content and SERPs during QRIs with useful documents.

Vakkari et al. [32] observed by eye-tracking how users' inspection of metadata in book pages was associated to novels' interest grading. The authors found that most time was used to assess a somewhat interesting novel compared to a not interesting or a very interesting novel. The findings show that the association of time use and novels' interest grading was non-linear [32].

Searching for books on the web and in digital libraries. Kim et al. [9] focused on web-based book search behavior. Transaction log analysis revealed some typical patterns of users' query reformulation. Most users followed up a new query by ending their session or issuing a completely different query. Minority of users limited the original search by adding a filtering condition. Ballard & Blaine [1] showed that when limiting capabilities were automatically displayed, users were more likely to refine their search compared to not displaying the limiting options.

Koolen et al. [10] compared the usefulness of professional metadata and user-generated content in determining books' topical relevance in social book search. Findings show that professional metadata was often not enough to determine books' topical relevance. Also, user-generated content was more effective for social book search compared to professional metadata.

Mckay et al. [15] examined the influence on an e-book presentation on the length of reader interactions with e-books. Inconsistencies or errors in a cover image, a blurb and a table of contents were found to be associated to an undesirable short time-span reading: flaws encouraged the user to loan the book for a closer inspection but ended up in rejection rapidly.

Browsing for fiction in digital libraries. In fiction retrieval browsing has been a common search strategy [e.g. 8, 20, 25].

Pejtersen [20] has identified three browsing search tactics which apply to fiction searching in online catalogs and match the purpose of this study. An analytical search tactic is used when readers wish to access books about some topic. A browsing tactic is applied in situations when readers have only a vague idea of what they would like to read. Search by analogy is generated when readers want something similar to a novel they have read.

Thudt et al. [30] have introduced the concept of targeted browsing, which is equivalent to Pejtersen's [20] analytical (topical) search. In targeted browsing the reader has some idea of the book he wants to find, such as the subject of the book. Similar to Pejtersen's [20] browsing tactic, the concept of open-ended browsing [30] refers to the situation where a reader is browsing for just interesting books. The Bohemian bookshelf [30], where browsing was enabled through information visualization offering various access points to the fiction collection, was found to support especially the open-ended browsing search tactic [30].

Previous studies [e.g. 14, 31] have revealed that most search interfaces to digital libraries support topical search in the form of querying. However, when browsing for fiction in an enriched online library catalog, it was found that searching without a query was associated to higher search success [18] in an enriched library catalog. Moreover, it was found by Oksanen & Vakkari [19] that effort devoted to search results instead of querying was associated to finding interesting novels in open-ended browsing situations.

This is due to the serendipitous nature of fiction searching: readers' needs for fiction are often ill-defined and best met by serendipitous discoveries when browsing [3]

Tang et al. [25] conducted a user study of aNobii, an online book sharing website that enables social browsing, which refers to the ability to re-orientate browsing by social navigational tools [21, 25]. Author browsing was found to be the most efficient in terms of number of items examined. Browsing similar and friends' bookshelves produced more novel items and serendipitous choices. Jiang [8] examined the users of a social library system and noticed that in addition to searching by using the internal search engine, catalog browsing, associative browsing and social browsing were popular search strategies applied by them. Searching was found to be the most adopted mode, while browsing appeared to be the most widespread mode.

In all, previous research has revealed that document viewing behavior seems to predict document usefulness. When searching for books in digital libraries, user-generated metadata, limiting options and interface elements are associated to user behavior. When searching for fiction in digital libraries, the adopting of browsing tactics seems to predict search success. We set out to investigate how these elements are associated to document viewing behavior during QRIs when browsing for fiction.

3. RESEARCH METHOD
3.1 Participants
80 participants with genuine fiction reading interest were recruited in public libraries, in fiction reading groups and in Finland's Open University's writing and literature classes. The Snowball sampling method and a newspaper advertisement were used. Participants were offered a movie ticket as compensation for participation.

Background variables (i.e. gender and experience in using online catalogs) did not significantly differ between control and test group. In both groups, 18 % of the participants were male and 82 % female. The age distribution of participants in the test catalog was significantly higher compared to the age distribution in the control catalog (p=.015). However, as there was no significant difference in participants' experience in using online catalogs between control and test groups, the higher age distribution in the test group is unlikely to affect the results greatly.

3.2 Fiction Search Tasks
Five search tasks were designed based on previous research [4, 20, 21, 30]. In two search tasks, the participants were given an author or a topic with which to begin the search process. The remaining three search tasks reflect the idea of individual information needs as the participants were asked to proceed according to their personal preferences. Search tasks formed reflect the following typical search tactics in fiction searching: known author/title search, topical search, open-ended browsing, search by analogy and searching without conducting a query.

The aim of this study was to examine fiction readers' search actions in *browsing situations*. Thus, the ***"known author task"*** was excluded from the analysis. Also, when conducting the user tests it was noticed that the tasks open ended browsing and searching without a query resembled each other. Thus, solely the tasks ***"topical search", "open ended browsing"*** and **"search by analogy"** were included in the analysis of this study. The browsing tasks were presented as follows: ***"Topical search"****: "Find three novels about upper class life in the 19th century."* ***"Open ended browsing"****: "Find three novels that interest you which you would like to read."* ***"Search by analogy"****: "Think of*

and mention one novel that you have read and found interesting recently. Now search for three similarly interesting novels as the one you mentioned."

3.3 The Catalogs Used

As a traditional baseline catalog, Satakirjastot -service was used (https://www.satakirjastot.fi). As an enriched test catalog the BookSampo -service was used (www.kirjasampo.fi) Satakirjastot (Sata) is the web service of the city libraries of the Satakunta region in Finland. The service consists of a library catalog and an information retrieval system from given databases. In Sata the metadata for fiction contains bibliographic information added with subject terms from the fiction thesaurus Kaunokki. Cover images and blurbs from recently published books are also available. In Sata, the search options are basic and advanced search. Figure 1 presents the advanced search option in Sata.
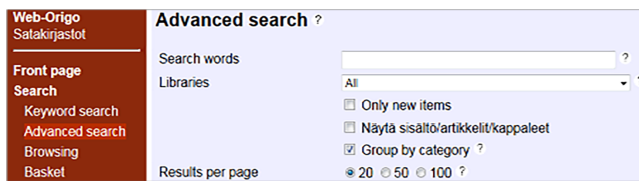

Figure 1. Advanced search in Satakirjastot.

In Sata, the search results are sorted according to format and relevance (such as book, DVD, CD). By choosing the preferred format the search results are presented in a new display according to author, title, publication year or class. Figure 2 presents the search result page in Sata.

Legend: 1= Author information, 2= Title, 3= Format (book)

Figure 2. Search result page in Satakirjastot.

Sampo includes metadata on the adult fiction collection from the HelMet Library which is the web service of the city libraries of the Helsinki Metropolitan area in Finland. Compared to the baseline catalog Sata, Sampo provides fiction metadata that is more varied and accounts for different access points to literature (Table 1). In Sampo, the associations and similarities between the works of literature are realized by semantic web technologies such as the ontologization of the fiction thesaurus Kaunokki, which is used for fiction indexing in Sampo [7].

Table 1. The major characteristics of the catalogs used.

Characteristics		Sampo	Sata
Searching	Quick search	Yes	Yes
	Adv. search	No	Yes
	Image search	Yes	No
	Browsing	Yes	Yes, after querying
Book Page Info	Author, title	Yes,	Yes, always
	Keywords	Yes, often	Yes, often
	Content	Yes, often	Yes, seldom
	Cover image	Yes, often	Yes, seldom

The front page of Sampo offers a variety of access points to the collection. Users can begin the search process with a basic search (similar to baseline catalog), by browsing the images of book covers, by skimming other users' virtual bookshelves or looking at the subject term cloud. The basic search option and a search result page in Sampo are presented in Figure 3.

Legend: 1=Basic search, 2= Results by author, 3= Results by genre, 4=Snippet

Figure 3. Basic search and search result page in BookSampo.

The search results are ranked by an author and the genre of literature (novels, plays, short stories) by matching keywords to metadata. The book page represents the content of a particular novel in detail. Book descriptions, automatic recommendations and text samples are also included.

When the catalogs for the experiment were considered, the aim was to identify a traditional catalog with a fiction collection comparable to Sampo in size and content. The collection in Sata was found to be relatively similar to Sampo's collection. The fiction collection of Sata consisted of 77 000 titles and the fiction collection of Sampo of 73 000 titles, both include novels, plays and poems. Both collections included the classics and recently published titles of Finnish and foreign literature. When the data was collected, e-books were not available in the fiction collections of Sata and Sampo.

Before the experiment, the researchers conducted several test queries in both catalogs in order to compare the results. They were found to share several overlapping titles. The analysis of Sata and Sampo revealed that the size and content of these collections were comparable, and therefore they could be used in the experiment.

3.4 Experimental Procedure

Before conducting the user tests, the experimental setting was pre-tested. The experiment consisted of the following steps: 1. Pre-questionnaire (demographic questions, participant's search experience in online catalogs and their reading interest). 2. Introduction to the experiment. 3. Demo of the retrieval system. 4. Execution of five search tasks. 5. Post-task questionnaire after each completed search task. 6. Post-session questionnaire after the completion of all search tasks. 7. Brief post-task interview.

The time for completing the tasks was not limited. The participants were randomized into control and test groups. Each participant completed the tasks either on Sampo or Sata. Latin square rotation was used with the tasks. Each participant

completed the tasks individually. During the experiment, the researcher was present to help in case technical problems occurred. The search logs were saved with Morae-software (http://www.techsmith.com/morae.html).

In each task, the participants were asked to search for three novels that were of interest to them. After each search task, the participants were asked to rank the novels selected according to how much they were of interest to them with an ordinal scale ranging from 1 to 3, where 1 was "a little interesting", 2 was "somewhat interesting" and 3 was "very interesting".

4. ANALYSIS
4.1 Measures

Success in search tasks was measured by the books' interest grading given by participants. Search was considered as "successful" if it resulted in a selection of a novel that was given an interest grading of "very interesting". Search was considered as "somewhat successful" if it resulted in a selection of a novel that was given an interest grading of "somewhat interesting" or "little interesting". Search was considered as "unsuccessful" if it resulted in no selections.

80 user tests were manually coded in order to identify different QRI types, and measure the search actions during the identified QRI types. In both catalogs three types of query reformulation intervals were identified: a) Successful QRI with very interesting novels selected (named as Successful QRI), b) Successful QRI with a little and somewhat interesting novels selected (named as Somewhat successful QRI) and c) Unsuccessful QRI with no interesting novels selected (named as Unsuccessful QRI). After identifying the QRI types, a coding scheme was developed for search actions. Each possible search action was identified in both catalogs by a single coder. The coding scheme was slightly different between the two catalogs used as they offer different search features (Table 2).

Table 2. The coding schemes for Sata and Sampo.

Search actions during QRI in both catalogs
Duration of QRI
Query time (total time on issuing queries in basic and advanced searching)
SERP time (total time on all search result pages)
Dwell time on book pages (total time on all book pages
Number of SERP visits
Number of opened book pages
Search actions during QRI solely in Sampo
Pivot browsing time (total time on all pivot browsing actions)
Number of pivot browsing actions

In Sata, options for submitting queries were basic and advanced search. In Sampo, solely the basic search option was provided. In Sampo, browsing the collection was possible without querying. The concept of *pivot browsing* [21] was applied to measure search actions that re-orientated browsing to follow features such as virtual bookshelves, cover images and automatic recommendations. In Sata, pivot browsing was not possible.

The search actions were measured by their frequency and duration. Each of the search tasks was coded individually. Time on issuing queries was not included in the total duration of QRI, as the aim was to investigate users' actions between two queries.

4.2 Statistical Analysis

The variables in the coding scheme (Table 2) were not normally distributed (Shapiro-Wilk=.95, p<0.05 in each variable). Shapiro-Wilkinson's test was used as there were fewer than 50 cases per catalog [19]. Friedman's test was used to test differences in search actions between QRI types within a task and between tasks. When comparing the search actions between catalogs, the variables were again screened for normality. Variables were not normally distributed (Kolmogorov-Smirnov=.95, p<0.05 in each variable). Kolmogorov-Smirnov's test was used as there were over 50 cases [24]. Mann-Whitney's non-parametric U-test was used to test differences in search actions between QRI types between catalogs.

The search actions in the open ended browsing and search by analogy tasks did not statistically differ from each other. Thus, for the economy of analysis, they were collapsed into one search task referred to here as *a browsing task*.

To analyze whether the search actions during QRIs were associated to books' interest grading, multinomial logistic regression analysis was applied. The aim was to identify the search actions that predicted the selection of very interesting novels compared to less interesting ones. Before conducting regression analysis, the data was re-coded using QRI as the unit of observation. Books' interest grading given by participants was used as a dependent variable in multinomial logistic regression analysis. The variables presented at Table 2 were used as independent variables. When constructing the regression models the method Enter was used to add independent variables into the models. Enter is a method where the researcher controls the entry of variables and the advancement of the regression process [24].

5. SEARCH ACTIONS DURING QRIs

In total, the data consisted of 1115 query reformulation intervals (Table 3). Only the valid sessions that contained at least two queries were concerned. The number of unsuccessful QRIs was significantly (p=.000) higher in Sata compared to Sampo.

Table 3. Number of successful, somewhat successful and unsuccessful QRIs in Sata and Sampo.

Measure	Sata n=626	Sampo n=489	p
# Successful QRIs	189	174	.974
# Somewhat successful QRIs	163	176	.752
# Unsuccessful QRIs	274	139	.000

5.1 Search Actions in Sata

In *"topical search"*, the successful and unsuccessful QRIs differed significantly from somewhat successful QRIs (Table 4). Somewhat successful QRIs were significantly longer with more time on SERPs and book pages compared to successful and unsuccessful QRIs. Time per an individual SERP and book page visits did not significantly differ between QRI types. Interestingly, more time was used on issuing queries before unsuccessful QRIs compared to successful and somewhat successful QRIs.

More SERPs were observed and book pages opened during somewhat successful QRIs compared to successful and unsuccessful QRIs. In unsuccessful QRIs book pages were examined significantly more infrequently and significantly more time was devoted to search results examination compared to successful and somewhat successful QRIs.

Table 4. Means (standard deviations) of search actions in QRIs in *"topical search"* in Sata. (Duration in seconds.)

Measure	Mean (st.dev)			Between QRI type p-values		
	Suc. QRI n=43	**Some. Suc. QRI** n=73	**Un-suc. QRI** n=91	Suc. Some. Suc.	Suc. Un-suc.	Some. Suc. Un-suc.
QRI duration	44 (54)	118 (113)	43 (63)	**.006**	.622	**.009**
Query time	16 (24)	20 (15)	61 (72)	.206	**.004**	.105
SERP time	27 (35)	58 (57)	34 (52)	.058	.411	**.004**
BP time	17 (23)	59 (65)	9 (16)	**.004**	.083	**.000**
# SERP visits	1.6 (1.9)	3.6 (3.4)	3.0 (2.9)	**.005**	.056	.602
# BPs	1.2 (1.6)	3.3 (3.4)	0.8 (1.5)	**.009**	.117	**.000**
Legend: Suc= Successful QRI, Some. Suc= Somewhat successful QRI, Unsuc=Unsuccessful QRI, BPs=Book pages						

In summary, in *"topical search"* the findings indicate that searchers needed most time for assessing the novel as somewhat interesting, and equally time for assessing the novel as very or not interesting. Searchers observed clearly more SERPs and book pages when assessing somewhat interesting novels compared to very interesting or not interesting ones. They tended to open several book pages before deciding the novel to be somewhat interesting. Consequently, SERP time and dwell time on book pages were the longest when somewhat interesting novels were selected. Interestingly, most time was devoted to querying before unsuccessful QRIs. This suggests, that before unsuccessful QRIs especially finding of correct query terms and conducting queries correctly was problematic to searchers. As a result, the issued queries did not produce desirable results from which to proceed to book pages and book pages were rarely opened.

In *"browsing task"*, the successful and somewhat successful QRIs differed significantly from unsuccessful QRIs (Table 5). Searchers invested significantly more time on SERPs and book pages, and opened more book pages during successful and somewhat successful QRIs compared to unsuccessful QRIs. In unsuccessful QRIs book pages were infrequently opened. Time per an individual SERP and book page visit did not significantly differ between QRI types. Like in *"topical search"*, most time was used on issuing queries before unsuccessful QRIs.

To summarize, in *"browsing task"* successful and somewhat successful QRIs resemble each other and differ from unsuccessful QRIs. In *"browsing task"* searchers need most time for assessing novels as very interesting, and least time for assessing novels as not interesting. Assessing somewhat interesting novels is commonly in the middle. Searchers observed clearly more SERPs and book pages when assessing interesting novels compared to not interesting ones. As a result, SERP time and dwell time on book pages were longer when interesting novels were chosen compared to not interesting ones. Similarly to *"topical search"*, searchers devoted most time on issuing queries before unsuccessful QRIs together with very few opened book pages. Again the finding

suggests that before unsuccessful QRIs, the issued queries do not produce desirable results from which to proceed to book pages. Therefore, book pages were rarely opened.

Table 5. Means (standard deviations) of search actions in QRIs in *"browsing task"* in Sata. (Duration in seconds.)

Measure	Mean (st.dev)			Between QRI type p-values		
	Suc. QRI n= 146	**Some. suc. QRI** n=90	**Un-suc. QRI** n=183	Suc. Some. suc.	Suc. Un-suc.	Some. suc. Un-suc.
QRI duration	132 (74)	81 (67)	57 (69)	.114	**.000**	**.027**
Query time	23 (20)	19 (30)	36 (40)	.150	**.000**	**.000**
SERP time	78 (44)	48 (36)	45 (59)	.114	**.000**	.206
BP time	54 (46)	38 (41)	12 (16)	.250	**.000**	**.000**
# SERP visits	3.1 (2.0)	2.6 (2.4)	3.1 (2.6)	.096	.505	.194
# BPs	2.7 (2.0)	2.2 (2.4)	0.9 (1.6)	.237	**.000**	**.002**
Legend: Suc= Successful QRI, Some. Suc= Somewhat successful QRI, Unsuc=Unsuccessful QRI, BPs=Book pages						

Differences in search actions between QRI types were examined also between tasks. Compared to *"browsing task"*, successful QRIs in *"topical search"* were significantly (p<.001) shorter. When very interesting novels were chosen significantly (p<.01) more SERPs were observed and more book pages were opened in browsing task compared to topical search. Therefore, SERP time and dwell time on book pages were significantly (p<.001) longer in *"browsing task"* compared to *"topical search".*

When assessing somewhat interesting novels, no significant differences occurred in search actions between tasks. When assessing not interesting novels in *"topical search"*, QRIs were significantly (p<.01) shorter with significantly less time on SERPs compared to *"browsing task"*.

5.2 Search Actions in Sampo

In *"topical search"* in Sampo, the successful and unsuccessful QRIs differed from somewhat successful QRIs (Table 6). Time devoted to SERPs and book pages was longer in somewhat successful QRIs compared to successful and unsuccessful QRIs. Time per an individual pivot browsing action, SERP and book page visits did not significantly differ between QRI types. Time devoted to issuing queries was substantially the same before each QRI type. Pivot browsing was scarce in each QRI type. Observed SERPs and opened book pages were more numerous in somewhat successful QRIs compared to successful and unsuccessful QRIs. In unsuccessful QRIs book pages were examined infrequently.

In all, in *"topical search"* searchers needed most time for assessing the novel as somewhat interesting and, almost equally, time for assessing the novel as very or not interesting. Searchers viewed visibly more SERPs and book pages when assessing somewhat interesting novels compared to very interesting or not interesting ones. They tended to open several book pages before

deciding the novel to be somewhat interesting. Consequently, SERP time and dwell time on book pages were the longest when somewhat interesting novels were selected. Querying or pivot browsing did not differ between QRI types.

Table 6. Means (standard deviations) of search actions in QRIs in *"topical search"* in Sampo. (Duration in seconds.)

Measure	Mean (st.dev)			Between QRI type p-values		
	Suc. QRI n=40	Some. suc. QRI n=74	Un-suc. QRI n=55	Suc. Some. suc.	Suc. Un-suc.	Some. suc. Un-suc.
QRI duration	68 (84)	105 (85)	60 (92)	**.027**	.724	**.001**
Query time	17 (27)	18 (19)	28 (49)	.114	.715	.516
Pivot time	3 (10)	3 (10)	6 (15)	.969	.206	.405
SERP time	37 (51)	55 (50)	39 (60)	.058	.480	**.023**
BP time	27 (40)	47 (46)	15 (35)	**.046**	**.033**	**.000**
# pivot actions	0.4 (1.0)	0.5 (1.2)	0.5 (1.3)	.905	.527	.763
# SERP visits	1.4 (1.7)	2.6 (2.2)	1.9 (3.3)	.052	.462	**.000**
# BPs	1.2 (1.9)	2.4 (2.7)	0.9 (2.4)	**.028**	**.050**	**.000**
Legend: Suc= Successful QRI, Some. Suc= Somewhat successful QRI, Unsuc=Unsuccessful QRI, BPs=Book pages						

In *"browsing task"*, the successful and somewhat successful QRIs differed from unsuccessful QRIs (Table 7). In successful and somewhat successful QRIs pivot browsing, viewed SERPs and opened book pages were more numerous compared to unsuccessful QRIs. As a result, SERP time and dwell time on book pages were significantly longer in successful and somewhat successful QRIs compared to unsuccessful QRIs. Time per an individual pivot browsing action, SERP and book page visit did not significantly differ between QRI types. Like in topical search, in unsuccessful QRIs book pages were infrequently opened.

To sum up, in *"browsing task"* successful and somewhat successful QRIs resembled each other and differed from unsuccessful QRIs. In *"browsing task"* searchers needed most time for assessing novels as very interesting, and least time for assessing novels as not interesting. Assessing somewhat interesting novels was commonly in the middle. Searchers observed visibly more SERPs and book pages when assessing interesting novels compared to not interesting novels. As a result, SERP time and dwell time on book pages were longer when interesting novels were chosen compared to when not interesting novels were selected.

Differences in search actions between QRI types were also examined between tasks. Compared to *"browsing task"*, successful QRIs in *"topical search"* were significantly (p<.001) shorter with less time on issuing queries and pivot browsing. Significantly (p<.05) more pivot browsing actions were made,

SERPs were viewed and book pages were opened in *"browsing task"* compared to *"topical search"* when very interesting novels were chosen. Searchers devoted significantly (p<.01) more time on SERPs and book pages when assessing a very interesting novel in *"browsing task"* compared to *"topical search"*.

Table 7. Means (standard deviations) of search actions in QRIs in *"browsing task"* in Sampo. (Duration in seconds.)

Measure	Mean (st.dev)			Between QRI type p-values		
	Suc. QRI n=134	Some. Suc. QRI n=102	Un-suc. QRI n=84	Suc. Some. suc.	Suc. Un-suc.	Some. Suc. Un-suc.
QRI duration	178 (140)	148 (115)	76 (89)	.118	**.000**	**.002**
Query time	21 (18)	15 (13)	27 (29)	**.002**	.806	.059
Pivot time	20 (37)	28 (41)	16 (27)	.317	.371	**.041**
SERP time	84 (71)	64 (64)	51 (64)	.071	**.008**	**.005**
BP time	74 (76)	56 (47)	9 (15)	.172	**.000**	**.000**
# pivot actions	1.1 (1.4)	1.4 (1.7)	0.5 (0.9)	.414	**.012**	.061
# SERP visits	2.8 (2.8)	2.0 (1.7)	1.7 (1.8)	**.032**	**.028**	**.050**
# BPs	2.6 (2.7)	2.1 (1.7)	0.5 (0.8)	.233	**.000**	**.001**
Legend: Suc= Successful QRI, Some. Suc= Somewhat successful QRI, Unsuc=Unsuccessful QRI, BPs=Book pages						

When assessing somewhat interesting novels, significantly (p<.001) more time was devoted to pivot browsing and more pivot browsing actions were conducted in *"browsing task"* compared to *"topical search"*. When selecting somewhat interesting novels, significantly (p<.05) more SERPs were viewed in *"topical search"* compared to *"browsing task"*. When assessing not interesting novels, significant differences in search actions during QRIs between tasks did not occur.

Like in Sata, identifying very interesting novels required more effort in *"browsing task"* compared to *"topical search"*. Moreover, when identifying interesting novels in Sampo pivot browsing was chosen as a search strategy more frequently in *"browsing task"* compared to *"topical search"*.

5.3. Summary of Findings

In both catalogs the major differences between tasks occurred in the number of observed SERPs and opened book pages during the QRI types. In both catalogs, in *"topical search"* very interesting novels were selected by viewing only one to two SERPs and opening only one book page. When selecting somewhat interesting novels, three to four SERPs were observed and three book pages opened, while assessing not interesting novels, book pages were rarely opened. In *"topical search"* selecting somewhat interesting novels required more effort compared to

selecting interesting or non-interesting novels, i.e. the association is curvilinear. In unsuccessful QRIs the inspecting of SERPs provided enough information to decide the results to be not interesting.

Diverging from the previous, in *"browsing task"* very interesting novels were selected by viewing three SERPs and opening three book pages. When assessing somewhat interesting novels, two SERPs were viewed and two book pages opened. When assessing not interesting novels, book pages were rarely opened. Thus, in *"browsing task"* the major differences in search actions between QRI types were linear: selecting interesting novels required most effort, while selecting not interesting ones required least effort.

6. MODELING SEARCH SUCCESS IN QRIs

The association of search actions during QRIs and books' interest grading was examined by multinomial logistic regression analysis.

Table 8. Multinomial regression models for *"topical search"*.

Independen. variables	Sata n=207			Sampo n=169		
	VI vs NI	SI vs NI	VI vs SI	VI vs NI	SI vs NI	VI vs SI
	Odds ratio					
Query time	.95 *	.90 ***	.95 *	.98	.96 **	.98
Pivot time	-	-	-	.98	.90 **	.92*
SERP time	1.04 **	1.05 **	1.00	1.02	1.02*	1.00
BP time	1.05	1.10 **	1.05 **	1.06 **	1.05*	.99
# SERP visits	.11 ***	.12 ***	1.13	.37 *	.27 **	.72
# BPs	6.08 **	3.73 *	.61	1.12	2.41*	2.05 *
Model fitting	χ^2 =151.574, df =10, p=.000, Nagel. Pseudo R-square = 0.59 Overall % predicted correctly 68 %			χ^2 = 58.654, df = 10, p=.000, Nagel. Pseudo R-square = 0.33 Overall % predicted correctly 65 %		

* = p < .05 ** = p < .01 *** = p < .001
Legend: VI=Very Interesting, SI=Somewhat Interesting, NI=Not Interesting, BPs=Book pages

In *"topical search",* the model for Sata suggests that when deciding between an interesting (VI and SI) and a not interesting novel (NI), the most significant predictors of the selection were the number of SERP visits and opened book pages (Table 8). An increase in the number of opened book pages produced six times greater odds to select a very interesting book compared to a non-interesting one, while a decrease in the number of viewed SERPs produced 9.1 (1/0.11) times greater odds of selecting a very interesting book compared to a non-interesting one. The association was similar in selecting between a somewhat interesting and a not interesting novel (SI and NI): the more book pages opened and the less SERPs viewed, the more likely the novel was assessed as a somewhat interesting compared to a not interesting. Interestingly, short query time and long dwell time on book pages contributed to predicting the selection of a very interesting novel instead of a somewhat interesting novel (VI and

SI). The total variance explained by the model was rather high for Sata, 59 % as indicated by Nagelkerke R^2.

In *"topical search"*, the model for Sampo suggests that when deciding whether the novel was very interesting or not interesting (VI and NI), the most significant predictor of the selection was the number of viewed SERPs. A decrease in the number of observed SERPs produced 2.7 (1/0.37) times greater odds of selecting a very interesting book compared to a non-interesting one. Like in Sata, when selecting between a somewhat interesting and a not interesting novel (SI and NI), an increase in the number of opened book pages produced two times greater odds to select a somewhat interesting book compared to a non-interesting one, while a decrease in the number of observed SERPs produced 3.7 (1/0.27) times greater odds of selecting a somewhat interesting book compared to a non-interesting one. Differing from Sata, when deciding either the novel was very interesting or somewhat interesting (VI and SI), the most significant predictor of the selection was the number of book pages opened. An increase in the number of opened book pages produced two times greater odds to select a somewhat interesting book compared to a very interesting one.

Table 9. Multinomial regression models for *"browsing task"*.

Indepen. variables	Sata n=419			Sampo n=320		
	VI vs NI	SI vs NI	VI vs SI	VI vs NI	SI vs NI	VI vs SI
	Odds ratio					
Query time	.94 ***	.95 **	1.01	.96 ***	.95 ***	.99
SERP time	1.04 ***	1.04 **	.99	1.00	1.00	1.00
BP time	1.04 *	1.04 *	.99	1.0 ***	1.07 **	.99
# SERP visits	.19 ***	.25 ***	1.31	.27 ***	.25 ***	.93
# BPs	3.50 **	3.27 **	.93	1.92	2.3 *	1.21
Model fitting	χ^2 =252.420, df =10, p=.000, Nagel. Pseudo R-square=0.52 Overall % predicted correctly 68 %			χ^2 = 148.036, df = 8, p=.000, Nagel. Pseudo R-square = 0.42 Overall % predicted correctly 55 %		

* = p < .05 ** = p < .01 *** = p < .001
Legend: VI=Very Interesting, SI=Somewhat Interesting, NI=Not Interesting, BPs=Book pages

In *"browsing task"*, the model for Sata suggests that when deciding whether the novel was interesting (VI and SI) or not interesting (NI), the most significant predictors were again the number of viewed SERPs and opened book pages (Table 9). An increase in the number of opened book pages produced almost four times greater odds to select a very interesting book compared to a non-interesting one, while a decrease in the number of viewed SERPs produced 5.3 (1/0.19) times greater odds of selecting a very interesting book compared to a non-interesting one. The trend was similar in selecting somewhat interesting books.

When selecting between very and somewhat interesting novels (VI and SI), the regression analysis yielded no significant

predictors. This is in line with findings in the previous section, as it was found that successful and somewhat successful QRIs in the browsing task resemble each other, and differ from unsuccessful QRIs.

In *"browsing task"*, the model for Sampo suggests that when selecting between very interesting and not interesting novels (VI and NI), the most significant predictor was the number of viewed SERPs. A decrease in the number of viewed SERPs produced 3.7 (1/0.27) times greater odds of selecting a very interesting book compared to a non-interesting one. The model also shows that quick queries and long dwell time on book pages predict the selection of very interesting novels compared to not interesting ones (VI and NI).

Similarly to *"topical search"*, the most important predictors of selecting a somewhat interesting novel instead of a not interesting one (SI and NI) were the number of viewed SERPs and opened book pages. An increase in the number of opened book pages produced two times greater odds to select a somewhat interesting book compared to a non-interesting one, while a decrease in the number of explored SERPs produced four (1/0.25) times greater odds of selecting a somewhat interesting book compared to a non-interesting one. Similarly to Sata, when deciding between a very interesting or a somewhat interesting novel (VI and SI), the regression analysis yielded no significant predictors of selection.

7. DISCUSSION

As far as we know, this was the first study to examine fiction readers' search behavior during query reformulation intervals. We compared readers' search actions during successful and unsuccessful QRIs in two browsing tasks and in two catalogs.

7.1 Search Actions between Tasks

Significant differences were found in search actions between browsing tasks. Our findings demonstrated that searching for topical novels differed significantly from explorative browsing for interesting novels. Topical search for novels appeared to share characteristics with searching for topically relevant non-fiction documents. Gwizdka [5] noticed that when searching for non-fiction documents, the highest cognitive effort was required to assess partially relevant documents compared to irrelevant and relevant documents. A partially relevant document was defined as a topically relevant document not containing the answer to a particular question. Similarly to [5], we found that when searching for topically relevant novels, most effort was required to assess somewhat interesting novels compared to very or non-interesting novels. In the topical search the duration of QRI and dwell time on book pages were the longest during somewhat successful QRIs and the shortest during unsuccessful QRIs, i.e. the association was curvilinear. Similarly to [32] it seems that when browsing for topical novels, in *"selection stage most time is used for judging borderline cases compared to more valuable or valueless cases"*.

Deviating from the topical search, our results showed that in explorative browsing for novels, the duration of QRI and dwell time on book pages were linearly associated to document value. Selecting interesting novels required most effort, while non-interesting ones least effort in viewing results and book pages, but most effort in querying. The finding is in line with Liu & Belkin [12] and Liu et al. [13] who found that when selecting non-fiction documents, dwell time on document pages was linearly associated to document usefulness. Differing from the topical search, it seems that when browsing for interesting novels, in the selection stage most time is used for judging the most interesting cases. The

difference could be partly explained by observing the participants: in the topical search they tended to select first a novel with high topical value and then assess its interest level. Thus, interest level was often a secondary criterion for selection in the topical search, whereas in the browsing task searcher's genuine interest was a strong motive to select a certain novel. It seems that identifying truly interesting novels in open ended browsing situations requires long dwell time on book pages, while eliminating not interesting ones requires least effort.

Finally, Liu et al. [11] found that during successful QRIs the users viewed more content pages and observed more SERPs, and spent longer time on them compared to unsuccessful QRIs. The findings of our study confirm, but also extend this result. Similarly to [11] our results demonstrated that the QRI type was associated to the duration of QRI. However, deviating from [11] our results showed that the search actions during successful and unsuccessful QRIs varied between browsing task type. In the topical search, the duration of QRI, SERP time and dwell time on book pages were the longest during somewhat successful QRIs, while shorter in both in successful QRIs. In the browsing task, the duration of QRI, SERP time and dwell time on book pages were the longest during successful QRIs and shortest during unsuccessful QRIs. To sum up, task type was clearly associated to search behavior during successful and somewhat successful QRIs.

7.2 Search Actions between Catalogs

We found that search actions during successful QRIs did not considerably differ between catalogs. Differences were found between catalogs in search actions during unsuccessful QRIs: in Sata query time and the number of viewed SERPs were significantly greater compared to Sampo. In Sata an unsuccessful search pattern was detected: particularly at the beginning of each task searchers tended to end up going in circles, where much time was devoted to selecting suitable entry terms together with a quick glance over the results before issuing a subsequent query. When glancing at the results, they were noticed unsatisfactory, and book pages were not entered. In consequence, visits in SERPs were numerous together with long query time. In Sampo a similar pattern was not discovered.

In all, it seems that in the traditional catalog more effort was invested in QRIs with no novels chosen compared to the enriched catalog. The enriched catalog result list contained rich information on novels' content such as a cover image and a snippet of a novel's blurb. Enriched result list might draw searchers into book pages for more detailed examination, and support the identification of interesting novels. Moreover, in book pages automatic recommendations are generated and searchers may proceed without entering a subsequent query. In the traditional catalog the result list provides titles and author names. If they are found not interesting, a necessity for proceeding is a re-formulated query. This may explain the difference in search actions during unsuccessful QRIs between the catalogs.

7.3 Search Actions and Search Success

We examined whether the search actions during QRIs were associated to books' interest grading in browsing tasks. We discovered that the search actions that contributed to predicting the selection of more interesting novels a) varied slightly between catalogs but b) were similar between tasks within a catalog.

Overall, our results demonstrated that in both tasks and catalogs the selection between interesting and not interesting novels was associated to few viewed SERPs and numerous opened book pages. Less viewed SERPs could be seen as a consequence of

fewer queries. Thus, the first query should produce a satisfactory result list, from which to proceed to glancing at several book pages. It could be assumed that the more book pages are viewed, the higher the probability of encountering a very interesting one.

In the topical search, a difference occurred between catalogs when selecting between a somewhat interesting and a very interesting novel. In Sata, the longer the dwell time on book pages, the more likely the novel was assessed as very interesting instead of somewhat interesting. This suggests that longer dwell time on book pages was linearly associated to greater success, and is in line with Liu & Belkin's work 12]. In Sampo, the more book pages were opened, the more likely the novel was assessed as very interesting instead of somewhat interesting. This differs from Liu & Belkin's work [12] as instead of long dwell time on book pages, higher document usefulness was detected when the volume of viewed book pages was high. The difference might partly be due to differences in metadata elements between catalogs: as noted previously, in Sampo the result list and book pages contain detailed information on novels' content. In Sata, solely bibliographical and publication information together with subject terms are provided on book pages. With novels published recently, a cover image and a blurb might be available, but in a separate window. Therefore, in Sata long dwell time on book pages might be seen almost as a necessity for understanding what a novel is about, while in Sampo quick visits to book pages provide detailed information on a novel's content.

In the browsing task, when selecting between very and somewhat interesting novels, the regression analysis yielded no significant predictors of selection. This is reasonable, as we found that successful and somewhat successful QRIs in the browsing task resembled each other, and differed from unsuccessful QRIs.

7.4 Limitations and Implications

Firstly, the sample was biased towards females and the highly educated. Previous studies [i.e. 23, 26] have shown gender to be associated to different reading interests of pleasure readers. For example, women are more likely to be heavy readers than males and they tend to read a greater variety of books than men [23, 26]. In our study women's expected wider literary knowledge might have supported them to identify interesting titles better compared to men in the topical search. Then again, men's likely narrower reading preferences might have supported them to select interesting novels within a more limited range of authors and genres compared to women in the browsing task. By implication, males might have completed the browsing task quicker and with fewer search actions compared to women.

Secondly, the participants conducted search tasks in an experimental setting. Thus, it is not reasonable to generalize the findings to the population at large. Instead, the transferability of findings could be highlighted. As the search tasks reflected the major browsing tactics for fiction retrieval, it could be assumed that the search actions of fiction readers in another setting would share some characteristics with our results.

A few suggestions on system design can be made based on the findings. For the first, the high number of unsuccessful QRIs in the traditional catalog Sata hints that the searchers had a limited understanding on how to formulate effective queries [3]. Features such as Did you mean...? or semantic autocompletion could support searchers to avoid zero hit situations. For the second, as noted by McKay et al. [15] and Koolen et al. [10], content presentation and metadata are essential factors in determining books' usefulness when searching for books in digital libraries. In

our study the enriched metadata in SERPs and book pages in Sampo might have supported the searchers to identify interesting topical novels more efficiently compared to Sata. In Sata, incomplete or missing metadata might have caused similar user behavior as in the work of McKay et al. [15]. Displaying solely the author and the title of a novel might have encouraged searchers to open topically relevant book pages for closer examination, and to devote much time to identifying the novel's topical relevance. If no additional metadata was provided, the novel might have been rejected or assessed as a somewhat interesting. Thus, as an implication for system design it could be acknowledged that enriched interface elements (such as a snippet, and a cover image) in SERPs might support readers to identify topical novels more quickly compared to the traditional result list that provides solely titles and author names.

For the third, in the browsing task selecting the most interesting novels required most effort. Similarly to [10], displaying the content of a novel with both professional metadata and user-generated content might support readers to assess novel's interest level more efficiently compared to providing solely professional metadata. Judging the relevance of a document with limited metadata information on a traditional catalog was found to be challenging for searchers [3]. The missing blurbs, cover images and book reviews in Sata might have affected the participants to reject novels that would have been of interest to them if enriched metadata had been available. Thus, providing a diverse range of metadata elements in title pages supports readers to judge the books. In addition, options for efficient browsing could enhance the finding of similarly interesting titles [16]. In our study identifying interesting titles by pivot browsing was often experienced as frustrating by the participants. Providing options for limited browsing, for example within a particular genre, could enhance the possibility of serendipitous discoveries.

8. CONCLUSION

We analyzed fiction readers' search actions during successful, somewhat successful and unsuccessful QRIs in browsing tasks.. Our contribution was especially in demonstrating that browsing task type was clearly associated to readers' document viewing behavior, i.e. the number of viewed SERPs, opened book pages and dwell time on book pages. In all, our study both confirmed, and extended previous findings. Browsing for topical novels resembled the retrieving of topically relevant non-fiction documents. The association of dwell time on book pages and document value was found to be non-linear: most effort was required to select somewhat interesting novels. When browsing for good books, the association of dwell time on book pages and document value was found to be linear: most effort in terms of dwell time was required to assess the most interesting cases. Therefore, if dwell time on book pages is considered as relevance feedback, it should be interpreted with understanding of the search situation, especially the task type.

9. REFERENCES

[1] Ballard, T. & Blaine, A. (2011). User search-limiting behavior in online catalogs. *New Library World*, 112, 5/6, 261-273.

[2] Bates, M. J. (1990). Where should the person stop and the information search interface start? *Inform Process Manag* 26, 5, 575-591.

[3] Blandford, A., Makri, S., Gow, J. et al. (2007). A library or just another information resource? A case study of users'

mental models of traditional and digital libraries. *J Am Soc Inf Sci Tec* 58, 3, 433-445.

[4] Goodall, D. (1989). *Browsing in the public libraries.* LISU Occasional paper No 1. Loughborough: Library and Information Statistics Unit.

[5] Gwizdka, J. (2014). Characterizing relevance with eye-tracking measures. In *Proceedings of IIiX'14.* ACM New York, NY, 58-67.

[6] Hu, X. & Kando, N. (2014) Evaluation of music search in casual-leisure situations. Presented at Searching4Fun workshop at IIiX2014. In *Proceedings of IIiX'14.* ACM New York, NY, 6.

[7] Hypén, K. & Mäkelä, E. (2011). An ideal model for an information system for fiction and its application: Kirjasampo and Semantic Web. *Library Review*, 60, 4, 279 – 292.

[8] Jiang, T. (2013). An exploratory study on social library system users' information seeking modes. *J Doc*, 69, 1, 6-26.

[9] Kim, J.Y., Feild, H. & Cartright, M. (2012). Understanding book search behavior on the web. In *Proceedings of CIKM'12.* ACM New York, NY, 744-753.

[10] Koolen, M., Kamps, J. & Kazai, G. (2012). Social book search: comparing topical relevance judgements and book suggestions for evaluation. In *Proceedings of CIKM'12.* ACM New York, NY, 185-194.

[11] Liu, C., Gwizdka, J. and Liu, J. (2010). Helping identify when users find useful documents: examination of Query Reformulation Intervals. In *Proceedings of IIiX'10.* ACM New York, NY, 215-224

[12] Liu, C. & Belkin, N.J. (2010). Personalizing information retrieval for multi-session tasks: the roles of task stage and task type. In *Proceedings of SIGIR'10.* ACM New York, NY, 26-33.

[13] Liu, C., Belkin, N.J. & Cole, M.J. (2012) Personalization of search results using interaction behaviors in search sessions. In *Proceedings of SIGIR'12.* ACM New York, NY, 205-214.

[14] Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM* 49, 4, 41-46.

[15] McKay, D., Buchanan, G., Vanderschantz, N. et al. (2012). Judging a book by its cover: interface elements that affect reader selection of eBooks. In *Proceedings of OzCHI'12.* ACM, New York, NY, 381-390.

[16] McKay, D., Smith, W. & Chang, S. (2014) Lend me some sugar: borrowing rates of neighbouring books as evidence for browsing. In *Proceedings of JCDL 2014.* ACM, New York, NY, 145-154.

[17] Mikkonen, A. & Vakkari, P. (2012). Readers' search strategies for accessing books in public libraries. In *Proceedings of IIiX'12.* ACM, New York, NY, 214-223.

[18] Mikkonen, A. & Vakkari, P. (2014). Finding fiction: search moves and success in two online catalogs. Submitted for publication.

[19] Oksanen, S. & Vakkari, P. (2012). Emphasis on Examining Results in Fiction Searches Contributes to Finding Good Novels. In *Proceedings of JCDL 2012.* ACM, New York, NY, 199-202.

[20] Pejtersen, A. M. (1989). *The Bookhouse: modelling user's needs and search strategies as a basis for system design.* Risø report M-2794. Roskilde Risø National Laboratory.

[21] Peters, I. (2011). Folksonomies, social tagging and information retrieval. In A. Foster & P. Rafferty (Eds.), *Innovations in information retrieval. Perspectives for theory and practice* (pp 85-116). London: Facet Publishing.

[22] Pöntinen, J. & Vakkari, P. (2013). Selecting fiction in library catalogs: A gaze-tracking study. In *Proceedings of TPDL'13.* LNCS 8092. Springer, 72-83.

[23] Ross, C.S. 2001. Making choices: What readers say about choosing books to read for pleasure. *The Acquisition Librarian* 13, 25, 5-21.

[24] Tabachnick, B. G. & Fidell, L. S. (2013). *Using multivariate statistics.* 6th ed. Boston: Allyn and Bacon.

[25] Tang, M-C., Sie, Y-J. & Ting, P-H. (2014). Evaluating books finding tools on social media: a case study of aNobii. *Inform Process Manag* 50, 54-68.

[26] Tepper, S. J. 2000. Fiction reading in America: Explaining the gender gap. *Poetics* 27, 4, 255-275.

[27] The Power of Our Network. Strategic Plan 2015-2018. (2014). National Library of The Netherlands. Retrieved March 23, 2015 from: https://www.kb.nl/sites/default/files/docs/strategicplan-2015-2018.pdf

[28] The Reading Agency. (2013). *Library facts.* Retrieved October 26, 2014 from: http://readingagency.org.uk/news/library-facts004/

[29] The State of America's Libraries. (2014). A report from the American Library Association. Retrieved March 23, 2015 from:http://www.ala.org/news/sites/ala.org.news/files/content/2014-State-of-Americas-Libraries-Report.pdf

[30] Thudt, A., Hinrichs, U. & Carpendale, S. (2012). The Bohemian Bookshelf: Supporting serendipitous book discoveries through information visualization. In *Proceedings of CHI'12.* ACM, New York, NY, 1461-1470.

[31] Toms, E. 2000. Serendipitous information retrieval. In *DELOS workshop on "Information seeking, searching and querying in digital libraries",* 17-20.

[32] Vakkari, P. & Luoma, A. & Pöntinen, J. (2014) . Book's interest grade and dwell time in metadata in selecting fiction. In *Proceedings of IIiX'14.* ACM New York, NY, 28-37.

[33] Wildemuth, B. M. (2004). The effects of domain knowledge on search tactic formulation. *J Am Soc Inf Sci Tec* 55, 3, 246-258.

[34] Wilson, M. L. & Elsweiler, D. (2010). Casual-leisure search: the exploratory search scenarios that break our current models. *HCIR'10.* New Brunswick, NJ, 28-31.

Readers' interest criteria in fiction book search in library catalogs

Anna Mikkonen and Pertti Vakkari
School of Information Sciences, University of Tampere, Tampere, Finland

Abstract

Purpose – The purpose of this paper is to investigate fiction readers' interest criteria when selecting novels in library catalogs for various search tasks.

Design/methodology/approach – The data of the book selection behavior from 80 genuine fiction readers were collected using recorded interviews and conversations. The data were qualitatively analyzed. Reuter's categorization of the components of aesthetic relevance has contributed to the construction of interest dimensions.

Findings – A five-dimension categorization of interest criteria is presented based on fiction readers' interpretations of the influential factors in fiction book selection in different search tasks. The findings revealed that readers apply the identified interest criteria in a flexible and multiphase way depending to the search task and the system used. The findings showed a context-related pattern in readers' fiction book selections. A combination of readers' search capacities, "behind the eyes" knowledge, affective factors and a well-functioning interaction with a system used results in a successful book selection.

Originality/value – A five-dimension categorization of adult fiction readers' interest criteria was created based on their search behaviors in library catalogs. The results provide a systematic step toward a comprehensive understanding of readers' fiction book selection in digital environments.

Keywords User studies, Reading, Books, Book selection, Fiction readers, Interest criteria, Library catalogs

Paper type Research paper

Introduction

There is some evidence (e.g. Ross, 2001; Saarinen and Vakkari, 2013) that book selection in physical libraries is influenced by a complex combination of previous reading experiences, elements from the books themselves, affective and personal factors and the informational needs of a person (Ross, 2001). In physical libraries, fiction readers have developed effective selecting strategies for fulfilling these multipart desires for recreational reading materials. In addition to known item searches, "looking around" at the shelves with the help of "behind the eyes" knowledge is constantly involved when considering a selection or a rejection of any particular book (Goodall, 1989; Ross, 2001).

Despite the knowledge on fiction readers' selection strategies in physical libraries, there are only a few studies focussing on users' book selection in library catalogs (e.g. McKay *et al.*, 2012; Reuter, 2007). Fiction e-collections are emerging in digital libraries at a rapid speed (The Reading Agency, 2013; The State of America's Libraries, 2014) and fiction is increasingly accessed, searched for and read on mobile devices, tablets and laptops (Buchanan *et al.*, 2015; Hayles, 2008; Miall and Dobson, 2006). A little evidence on fiction readers' book selection in various digital environments is available to support the designing of interfaces for fiction book search. It is not known what kind of selection criteria fiction readers apply when choosing interesting titles, and to what extent current online library catalogs support the book selection process.

The notion of relevance defined by Cooper (1971) as "whether a piece of information is on a subject which has some topical bearing on the information need" functions ill-suited for the selection of recreational reading materials where the interaction with an information system often occurs without a topical search objective and where emotions and serendipity play a major role (Agosti *et al.*, 2014; Ross, 2001). Alternative evaluation criteria have been employed to measure untargeted searching behavior. For example, the concept of aesthetic relevance (Reuter, 2007) has been presented. However, a little attention has been given to the actual selection criteria applied by fiction readers in library catalogs. A potential means of tackling this issue is to let the readers themselves describe the influential factors in their book selection process and to apply these criteria in examining the book selection behavior.

In this paper, the interest criteria for novels are investigated as described by 80 genuine fiction readers while selecting novels in two library catalogs in various search tasks. The contributions of this paper are twofold. First, a five-dimension categorization of fiction readers' interest criteria for examining the book selection process is presented; second, an in-depth study of the fiction readers' interpretations of the influential factors in fiction book selection in different search tasks and library catalogs is provided.

The study seeks to address four research questions:

RQ1. What kind of interest criteria readers apply in selecting novels in fiction search?

RQ2. Do the readers' interest criteria vary between different search tasks? If yes, how?

RQ3. Do the applied criteria vary between a traditional library catalog and an enriched library catalog? If yes, how?

RQ4. Are there differences in the applied interest criteria between rejected and selected novels?

The study offers vital information on fiction readers' book selection behavior that could be used in designing interfaces for fiction. The paper presents a categorization of interest criteria as defined by the users themselves for evaluating performance in the fiction search. It may contribute to the designing of future studies for evaluating user behavior in digital libraries.

Literature review
Book selection in physical libraries
Previous research (e.g. Ooi and Liew, 2011; Ross, 2001; Saarinen and Vakkari, 2013) has identified several factors associated with fiction book selection in physical libraries. They include previous literary knowledge, literary preferences, mood and clues in the books themselves. Ross (2001) found that when choosing books to read for pleasure, readers used a variety of considerations to identify interesting novels. Ross (2001) refers to these considerations as "behind the eyes" knowledge where previous experience and meta-knowledge of authors, publishers, cover-art and recommendations from family or friends were involved when considering between selection and rejection of an item.

In a qualitative study of 16 fiction readers' book selection behavior, Saarinen and Vakkari (2013) found that in a quest for interesting novels on the shelves in a public library, fiction readers actively searched for clues of good novels in the books themselves. The authors suggested that as in Goodall (1989), the indicators of good

novels could be categorized into clues that trigger interest toward a novel; and clues that determine the decision of borrowing the novel. The most important indicators of good novels were the author's name, back cover text and scanning a particular item.

Ooi and Liew (2011) found that fiction readers tended to have a book title or author already in mind when they visited their public library. Thus, fiction selection was often based on a pre-determined decision. Instead of browsing in public libraries, participants actively sought recommendations for interesting books outside libraries. Browsing of fiction was usually focussed on the "New Books" display and the "Book returns" section (Ooi and Liew, 2011). Similarly, Saarinen and Vakkari (2013) found that the shelf for returned books was a common place for avid readers to begin their search.

Book selection in library catalogs

Previous research (i.e. Mikkonen and Vakkari, 2012; Pejtersen, 1989; Spiller, 1980; Yu and O'Brien, 1996) has revealed two major approaches for selecting novels: selecting known titles or authors and browsing for possibly interesting items without a clearly defined idea. When searching for specific books, the selection process rests greatly upon authors' names (Yu and O'Brien, 1996). Adkins and Bossaller (2007) compared different entry points to fiction collections in online bookstores, reader advisory databases and public library catalogs. Known item search was the most supported selection strategy by library catalogs. Compared with known item search, browsing is a more complex search strategy where user's literary competence, affective factors, clues in the books themselves and expectations toward reading material influence the selection process (Ross, 2001).

In selecting novels by browsing, the process might be guided by a particular topic or a previously read interesting title or it might be untargeted aiming to find just interesting books (Pejtersen, 1989; Thudt *et al.*, 2012). When browsing for fiction in online library catalogs, it has been found that searching without a query, effort devoted to search results instead of querying and examining of the author and title information instead of the content description is associated to successful book selection (Mikkonen and Vakkari, 2016; Oksanen and Vakkari, 2012; Pöntinen and Vakkari, 2013; Vakkari and Pöntinen, 2015).

Relevance in fiction book selection

The concept of relevance has been acknowledged as central to the theory of information retrieval and a fundamental concern in evaluating information retrieval systems (Borlund, 2003; Cooper, 1971). The studies on relevance have focussed either on a system-centered perspective on relevance as a logical and topical relationship between a user's query and a subject of a document, or on a user-centered approach (Borlund, 2003; Cooper, 1971; Saracevic, 1996). Saracevic (1996) distinguishes between five basic types of relevance: algorithmic, topical, cognitive, situational and affective relevance. Saracevic's affective relevance as the relation between the intents, goals and motivations of the user and the document fits to the emotion-bound fiction selection process (Ross, 2001). However, affective relevance in fiction book selection has not been examined.

Previous research on relevance assessments in fiction book selection in library catalogs has been conducted by Reuter (2007) and Koolen *et al.* (2015). Reuter (2007) examined children's book selection in a digital library with the concept of "aesthetic relevance." Aesthetic relevance was defined as the potential of a document to provide a

suitable reading experience. The most influential factors in children's book selections were the metadata (such as title, author), the extrinsic appeal of a book (such as recency, format) and accessibility (such as length, text density).

Koolen *et al.* (2015) investigated the relevance aspects expressed in book requests in the Library Thing discussion forums. As in Reuter (2007), the authors identified relevance aspects such as accessibility, content, engagement, familiarity, metadata, novelty, known item and sociocultural. The most adopted relevance aspects in book requests were the content of a book and looking for familiar reading experiences. Also, a unique combination of content, context and examples in book requests was detected.

Method

Participants

In total, 80 people with fiction reading interest were recruited in public libraries, in fiction reading groups and in writing and literature classes in the Open University of Finland. In addition, the Snowball sampling method and a newspaper advertisement were used. Participants were offered a movie ticket to participate in the study. Participants with no genuine fiction reading interest were excluded from the study.

Participants were randomized into control and test groups. In both the control and test group, 18 percent of the participants were male and 82 percent female. In both groups, the age distribution of participants varied from 20 to 80. In a traditional catalog participants averaged 34 years of age (SD 12.7). In an enriched catalog participants averaged 42 years of age (SD 16.8). In both catalogs, 18 percent of the participants had a middle-level education and 82 percent a high-level education. For a detailed description of the participants, see Mikkonen and Vakkari (2015).

Fiction reading activity and fiction reading preferences did not significantly differ between the control and test group. On average, the participants in Sampo read 26 novels a year while the respective figure in Sata was 18 ($t = 1.409$, $p > 0.05$). Participants were asked to fill in a pre-questionnaire measuring their reading preferences. The questionnaire was designed to measure motives for fiction reading and important features in fiction texts while reading fiction books. No significant differences occurred in the 38 item scale: participants valued entertainment, aesthetic experiences and utilitarian aspects of fiction reading similarly in both groups.

Search tasks

Four search tasks were designed based on previous research (Adkins and Bossaller, 2007; Goodall, 1989; Peters, 2011; Pejtersen, 1989; Ross, 2001; Spiller, 1980; Thudt *et al.*, 2012; Yu and O'Brien, 1996). In two search tasks, the participants were given an author or a topic with which to begin the search process. The remaining two search tasks reflect the idea of individual and dynamic information needs as the participants were asked to proceed according to their personal preferences without any given topic. Simulated search tasks were as follows:

Known author search: "A friend of yours recommends you to familiarize yourself with the novels of Olli Jalonen. Find Olli Jalonen's novels and choose two novels which are of interest to you."

Topical search: "Find three novels of interest about upper class life in the 19th century."

Open ended browsing: "Find three novels that interest you which you would like to read."

Search by analogy: "Think of and mention one novel that you have read and found interesting recently. Now search for three novels that you would consider similarly interesting as the one you mentioned."

The catalogs used

As a traditional catalog, the Satakirjastot-service was used (www.satakirjastot.fi). As an enriched catalog, the BookSampo-service was used (www.kirjasampo.fi). The concept of an enriched catalog (or a metadata-enriched library catalog) refers to a unique metadata that enriches the bibliographic description of a collection (such as book cover images, book descriptions, virtual book shelves) and enables users to browse and interact with the collection allowing them to read, add reviews and generate content in addition to professional created metadata. In recent years, it has become increasingly available in online library catalogs (Eden, 2002).

Satakirjastot (Sata) is the web service of the city libraries of the Satakunta region in Finland. The service consists of a library catalog and an information retrieval system for the given databases. In Sata, the metadata for fiction contains bibliographic information added with subject terms from the fiction thesaurus Kaunokki. Cover images and blurbs from recently published books are also available. In Sata, the search options are basic and advanced search.

BookSampo (Sampo) is an enriched web-service for fiction in Finland. In Sampo, the associations and similarities between the works of literature are realized by semantic web technologies such as the ontologization of the fiction thesaurus Kaunokki, which is used for fiction indexing in Sampo (Hypén and Mäkelä, 2011). The front page of Sampo offers a variety of access points to the collection. The book page represents the content of a particular novel in detail. Book descriptions, automatic recommendations and text samples are also included. For a detailed description of the catalogs used, see Mikkonen and Vakkari (2015).

Experimental procedure

Before conducting the user tests, the experimental setting was pre-tested with one participant. The experiment consisted of the following steps:

(1) pre-questionnaire including demographic questions, participants' search experience in online catalogs in general, participants familiarity with the catalog used in the experiment and their reading interest;

(2) introduction to the experiment;

(3) brief demo of the retrieval system (approximately two minutes);

(4) execution of four search tasks;

(5) a post-task questionnaire after each completed search task;

(6) a post-session questionnaire after the completion of all search tasks; and

(7) brief post-task interview after one search task (decided in advance).

The participants were randomized into control and test groups. Each participant completed the tasks either on Sampo or Sata. The pre-questionnaire yielded that the participants in both groups were unfamiliar with the catalogs used in the experiment. The known author task functioned as a training task and was conducted first with each participant. Latin square rotation was used with three other tasks.

Each participant completed the tasks individually. Public libraries, meeting rooms at a university, coffee shops and participants' homes were chosen as test environments according to participants' wishes. The time for completing the tasks was not limited. During the experiment, the researcher was present to help in case technical problems occurred. At the beginning of the search process, participants were not asked to follow the think-aloud strategy. As the experimental procedure contained various questionnaires and complex search tasks, it was considered as too stressful for the participants to say out loud what they were thinking. However, during the search process, a majority of the participants began naturally to "think-aloud" and give comments about their decisions while completing the search tasks. The researcher also asked questions on particular book choices during the search process. However, the researcher did not guide or help the participants in questions concerning the completion of the search tasks.

In each task, the participants were asked to search for three novels that were of interest to them. After each search task, the participants were asked to rank the novels found according to how much they were of interest to them with an ordinal scale ranging from 1 to 3, where 1 was "a little interesting," 2 was "somewhat interesting" and 3 was "very interesting." Scoring 0 was used if an interesting novel was not found.

Data and analysis
The search logs and participants' speech were recorded with Morae-software (www. techsmith.com/morae.html). For the purposes of this study, 80 participants' interviews and conversations with the researcher were transcribed. The audio material was transcribed word for word and qualitatively analyzed. The purpose of the qualitative data analysis was to identify the categories for the interest criteria for novels mentioned by the participants, and apply these criteria for examining participants' book selection behavior. To guide the qualitative data analysis, findings from previous research (Reuter, 2007; Ross, 2001; Saarinen and Vakkari, 2013) were taken into consideration. Particularly, the coding scheme for children's' book selection in a digital library developed and applied by Reuter (2007) supported the design of an initial coding scheme for this study.

Based on a preliminary review of the transcripts, the authors developed a coding scheme representing the factors that influenced fiction readers' book selection at each search task. The authors discussed and refined the code labels and definitions several times in order to ensure the coding scheme to be unambiguous with no overlapping sub-categories. A protocol for marking the utterances was then developed. An utterance was defined as a sentence or a unit of conversation covering a single aspect of a book selection. Utterance units that covered multiple topics were divided and each unit was given an appropriate code. Each utterance was given a single code, no multiple codes for a single utterance were allowed. The transcripts were then reviewed line by line and utterances were encoded according to the coding scheme. The final coding scheme with code definitions and example utterances is included in Appendix.

The total number of selection mentions in each dimension was first calculated over search tasks and catalogs. In addition, the total mentions concerning rejecting a book were calculated for each dimension. Then, the mentions in each dimension were calculated by tasks and by catalogs. To test for significant differences in the distribution of selection mentions between catalogs, χ^2 p-values were calculated.

Results
Readers' interest criteria for novels
The qualitative data contained a total of 931 selection mentions for novels, 607 from Sampo and 324 from Sata. Rejecting mentions were far less as the data included a total of 118 rejecting ones, 85 of them in Sampo and 33 of them in Sata. Table I presents the distribution of the selecting and rejecting mentions over five major dimensions for novels' interest criteria.

The results show that when selecting novels in two library catalogs, the most applied interest criteria by readers were familiarity and bibliographic information. Familiarity was related to known item search, as the majority of the selecting mentions in familiarity covered selecting novels from a known author or searching for a known title. It seems that a common way of searching for interesting novels in library catalogs is to select books based on previous reading experiences, to select well-known authors and to select novels based on recommendations from the media.

When selecting a novel based on bibliographic information, the most important interest criterion was the title of a novel. In each task and in both catalogs, the title received a clear majority of the selecting mentions. In addition, the publication date and format of a book were also often mentioned. Readers' common habit was to select books published recently instead of older books. The format of a book was associated to a short story as a form of literature, as the mention of "short story" in bibliographic information either caused an immediate rejection or a book selection with delight.

The content dimension included selecting mentions related to the intellectual content description provided by library professionals such as subject headings, description of a novel's plot (blurb) and genre classification. Perhaps surprisingly, the content description received fewer mentions among readers' interest criteria compared with the bibliographic information. Content description was emphasized more over bibliographic information solely in the topical search task, which resembled a non-fiction search. This hints that when selecting topical novels, the content description influences the selection process as topically non-relevant novels are easily eliminated by examining the subject headings and the blurb. When selecting interesting titles according to one's genuine reading preferences, familiarity and bibliographic information seem to have greater popularity over the content description.

The engagement in reading as an interest criterion refers to the mentions indicating general liking or disliking of a novel. Readers often reflected the ideas and expectations triggered by a particular novel and made the selections based on these preliminary estimations. Mentions such as "This sounds really interesting" and "I would choose this book for real" were common in the engagement dimension.

Sociocultural aspects were the least mentioned in readers' book selections compared with the other major dimensions. Most of the selecting mentions in sociocultural

	Dimension	Selecting mentions ($n = 931$)	Rejecting mentions ($n = 118$)
Table I. Distribution of selecting and rejecting mentions by the interest criteria dimensions (percent)	Familiarity	30	9
	Bibliographic information	30	43
	Content	21	44
	Engagement	13	4
	Sociocultural	6	0
	Total	100	100

dimension were related to identifying a personal connection to a particular novel which triggered an interest to select the book. The personal connection might have been identified from a title of the novel or by examining the blurb and the subject headings on a title page:

> Becoming father and daughter. This sounds interesting as I have only daughters, not sons, and the relations of fathers and daughters are always interesting (KSk11).

In both catalogs, the rejecting mentions were most often associated to bibliographic information such as an author in a sense that readers had a clear vision of the undesired authors. Also, publication date was mentioned as readers often avoided older books. The rejecting mentions were also associated to a single subject term on a title page: an undesired theme or topic was easily detected from the subject headings section and the novel was quickly rejected.

Interest criteria in different search tasks and catalogs

Known author search. The findings show that when the author was given, readers selected the novels mostly based on bibliographic information, content description and with previous knowledge on the given author's literary production (Table II).

The major difference compared with the other search tasks was the notable influence of a novel's title in the book selection in the known author search. In Sata, 42 percent of the mentions concerned the title while the respective figure in Sampo was 24 percent. When the search results were examined, the titles were not just passively skimmed through. Instead, the participants analyzed and interpreted the ideas provoked by the titles in detail and evaluated the possible reading experience offered by a particular title. An appealing title was often mentioned as containing an aspect close to one's personal life. Also, the title could have been mentioned as engaging if it was funny, included interesting concepts or a clever play on words.

Overall the results show that when selecting books from a known author, bibliographic information turned out to contribute more to the book selection compared with the content description. The content description was often used as a secondary criterion for the selection after detecting an appealing title from the search results. In addition to the title, the publication year (Sampo 3 percent, Sata 11 percent) and genre of literature (Sampo 8 percent, Sata 3 percent) received some mentions. Especially in Sata, the publication year was an important criterion as participants favored novels published recently over older titles. The emphasis on bibliographic information over content description is an interesting notion: as the production of a given author was unfamiliar to the majority of the participants, it could have been expected that

Dimension	Selecting mentions	
	Sampo $n=100$	Sata $n=38$
Familiarity	21	11
Bibliographic information	39	55
Content	24	32
Engagement	8	0
Sociocultural	8	2
Total	100	100
χ^2 p-value	0.036	

Table II.
Selecting mentions
in the known
author search by
catalogs (percent)

especially the plot and the theme of a novel would have been more popular in book selection compared with bibliographic information. However, according to the data, this does not seem to be the case.

The findings show that familiarity and previous reading experiences as criteria were more important to the users of Sampo compared with the users of Sata. The post-task question yielded that the users of Sampo were significantly (t-test 2.326, $p = 0.023$) more familiar with the given author's literary production compared with the users of Sata. This explains the difference in the selecting mentions related to familiarity between catalogs.

Topical search. Table III shows that when the topic was given, readers selected the novels mostly based on familiarity and the content description.

In topical search, familiarity with the given topic was a major factor in the book selection. In Sata, previous knowledge of suitable authors and titles was the most often mentioned single criterion when selecting topical novels by the share of 39 percent. The respective figure in Sampo was 25 percent. Participants used their knowledge of authors and titles in two ways when selecting topical novels: they issued a keyword search with a suitable author or title; or they issued a keyword search with a topic (e.g. upper class nineteenth century) and selected known items from the result list. The post-task question showed no significant difference (t-test 0.674, $p = 0.503$) in participants' familiarity with the given topic between the users of Sata and Sampo.

Compared with the other search tasks, topical search was the only one where content description overcame bibliographic information in selection. Both in Sampo and Sata, readers detected the topical relevance of a novel from the subject headings and the blurb on an item page. This explains why content description influenced the book selections greatly in the topical search. The results show that the enriched content description in the search results and item pages was the dominant interest criterion over the familiarity in Sampo. In Sata, the participants relied more on their knowledge of literature over examining the content description which was often limited and narrower than the one in Sampo. This might explain why the users of Sata emphasized known authors and titles as influencing their book selections more over Sampo.

The results also show that bibliographic information had a great role in selecting topical novels in both catalogs. The overwhelming majority of those mentions were about the title of a novel. Differing from the other search tasks, in topical search title mentions did not refer to particularly appealing or likable titles. Instead, they focussed on identifying classics such as *Jane Eyre* or *Pride and Prejudice* representing upper class life in the nineteenth century.

	Selecting mentions	
Dimension	Sampo $n = 99$	Sata $n = 64$
Familiarity	26	42
Bibliographic information	21	20
Content	36	30
Engagement	13	6
Sociocultural	4	2
Total	100	100
χ^2 p-value	0.063	

Table III.
Selecting mentions in the topical search by catalogs (percent)

Open ended browsing. In the open ended browsing task, there were no differences in selection criteria between catalogs. Readers in both catalogs relied greatly on their literary knowledge and familiarity with preferred authors and titles (Table IV).

Over half of the total selecting mentions in Sata, and over 40 percent in Sampo were related to previous reading experiences and items known from the media. A common pattern to select "just good books" was to search for favorite authors' novelties, or titles from personal reading lists kept and updated constantly on mobile devices and personal organizers. Selecting novels based on familiarity was detected at the beginning of the search task: if a participant was unable to come up with an author or a title after reading the indicative request, the task was experienced as challenging and difficult:

> This is very difficult. Not to have any hint or direction where to go or begin with (KSh8).

> I always know something about the writer or the book I want to discover. I never search for books this way (KSx23).

The lack of ideas explains why recommendations from the media received selecting mentions (8 percent in both catalogs) in the open ended browsing task more than in the other search tasks. Literary awards, television shows, newspaper articles and radio programs concerning recently published fiction books were mentioned as influencing the book selections in the open ended browsing task as they provoked ideas for the keyword searches.

Since participants' literary knowledge functioned as the most frequent interest criterion in the book selections for "just good reading," less emphasis was given to bibliographic information and the content description. When mentions concerned bibliographic information, the title of a novel was again the most important criterion (Sampo 15 percent, Sata 13 percent). Engagement in reading, especially a positive reading experience provoked by a novel, was also perceived as an important selection criterion in the open ended browsing task. This differs from the known author search and the topical search tasks and hints that when readers are selecting novels based on their true reading interests, the reading experience provided by a particular title is valued more compared to selecting topical novels or novels of a given author. It seems that when browsing for good reading in library catalogs, after having selected the known titles, the primary condition to select an item is an appealing title that provokes positive reading expectations and likely engages one in the reading process.

Search by analogy. In the search by analogy task, the participants selected three similarly interesting novels to a previously read one. At the beginning of the search task, the participants interpreted the idea of similarity and applied this interpretation to

Dimension	Selecting mentions	
	Sampo $n = 172$	Sata $n = 99$
Familiarity	43	55
Bibliographic information	18	16
Content	12	11
Engagement	17	11
Sociocultural	9	7
Total	100	100
χ^2 p-value	0.367	

Table IV.
Selecting mentions in
the open ended
browsing by
catalogs (percent)

their book selection process. Participants mentioned the similarity to be related to the following aspects in novels: the same genre, the same author, a similar topic, the same mood and a similar narrative style.

Depending on the aspect of similarity, the search task was perceived either as an easy or a challenging one. If similar novels were selected based on a genre or an author, the search task was often completed easily. If a similar topic, mood or a narrative style was the starting point for the selection, the participants had difficulties in discovering similarly interesting novels. The similarity was often mentioned being a combination of various similarity aspects, such as a novel with a particular narrative style covering similar topics. These expectations were rarely met and participants compromised between similar topics and interestedness.

Table V shows that in the search by analogy task, the most important criterion for the book selections in both catalogs was bibliographic information, particularly the title and the author of a novel.

In Sata, almost 32 percent of the mentions concerned the title, while in Sampo the respective figure was 21 percent. In both catalogs, approximately 6 percent of the mentions concerned the author information. Selecting similar novels based on an author or a title often resembled a topical keyword search: participants issued a query with the wished for topic of a novel, and selected novels from the search result list based on the author information and the appealing title.

The content description and familiarity with authors and titles in selecting similar novels was emphasized slightly more in Sampo than in Sata. Similarly, the users of Sampo mentioned engagement in reading notably more compared to the users of Sata. Almost 20 percent of the mentions in Sampo were based on general interest or the mood of a novel, or selecting books based on serendipity. The respective figure in Sata was 11 percent. It seems that in a traditional catalog, selecting similar novels was associated to detecting an appealing title after topical search, whereas in the enriched catalog the title, the content and the positive expectations provoked by the former were emphasized evenly. This might hint that the enriched content description, the various visual and social navigational tools and the automatic recommendations in Sampo engage the participants in the book selection process more than the query-based selection process in Sata. As the selection process turns into an amusing and engaging flow instead of designing one query after another, the reading materials selected might begin to fulfill readers' expectations often unknown or unconscious:

I didn't actually know what I wanted to find until this book came up. Now I realized this is what I am aiming for (KSy38).

	Selecting mentions	
Dimension	Sampo $n = 236$	Sata $n = 123$
Familiarity	23	19
Bibliographic information	30	43
Content	23	18
Engagement	20	11
Sociocultural	4	9
Total	100	100
χ^2 p-value	0.049	

Table V.
Selecting mentions in the search by analogy by catalogs (percent)

Selection and rejection of a novel
The findings from the four search tasks hint that in library catalogs, the process resulting in a book selection differs from the one resulting in a book rejection. Overall, it seems that the rejection of a book is a simple and quick process in which a single metadata element (such as a subject term, a publication year) often determines the decision of a non-interesting book and a rejection (see Table I). Quite the contrary, a book selection seems to be a complex chain of search actions, during which the influence of previous knowledge, expectations, emotions and interpretations is associated to the outcome. In library catalogs, a successful book selection combines successful information retrieval, reader's search skills, a strong capacity to apply one's knowledge on preferred literary genres, authors and topics to the book selection process, capacity to identify appealing items from a large amount of available readings, the interpretation of the expectations provoked by a particular title and system's support (such as a well-functioning search engine, complete and enriched metadata, a good recommender system). The challenge is to articulate one's idea of good reading as it is often unknown, especially at the beginning of the search.

The findings suggest that novels' interest criteria vary also within successful selection processes. According to a search task, readers create expectations toward the reading material to be selected and apply the interest criteria alternately to meet these expectations. The interest criteria for novels identified in this study seem to be well applicable to various search tasks for recreational readings, and readers seem to be skilled in altering the emphasis between the interest criteria in a dynamic way according to the search task.

Discussion
The primary motivation for undertaking this research was to understand fiction readers' book selection process in library catalogs and to identify the criteria for selecting a particular item in various search tasks. In this section, the findings are discussed in depth.

Influence on previous knowledge and metadata on book selection
Overall, the results of this study both confirmed and extended previous knowledge on fiction book selection. The interest criteria for novels as described by genuine fiction readers were found to bear similarity to the relevance aspects identified by Koolen *et al.* (2015) and Reuter (2007). Even though previous research had identified familiarity to be an influential factor in fiction book selection, readers' literary knowledge, literary preferences and a capacity to come up with authors and titles were found to have a remarkable role in the book selection process, particularly in browsing tasks. As the common selection strategies in public libraries (such as browsing the shelves of returned loans and new items) were not applicable in digital environments, readers were to discover alternative ideas for their book searches.

The book selection process was usually initiated by figuring out ideas for keyword searches such as known authors, titles or books nominated for a literary award. This confirms that searching for books based on authors' names (Spiller, 1980; Yu and O'Brien, 1996) applies also in library catalogs. At the beginning of a search task, a failure in discovering ideas for interesting reading based on literary knowledge turned out to be a major obstacle in the book selection process. The examination of participants' search behaviors revealed that a combination of good search skills,

knowledge of authors and titles and a positive association derived from metadata of a particular item resulted in the most successful book selections in both catalogs. Compared with Ross's (2001) findings of the book selection process in physical environments, it seems to be even more challenging in library catalogs. "Behind the eyes knowledge" offers an effective starting point for the search but alone it is not sufficient for a successful book selection.

As in Reuter (2007), the participants of this study emphasized the role of a novel's title in the book selection. The attributes for a good title were found to be context related. In searches for truly interesting readings, the titles were carefully examined and clues to personal links to a subject of a novel were sought. Preliminary reading expectations were commonly produced based on title names in the search results. In searches for topical novels, the title names were examined in a quest for well-known classics about the topic instead of interestedness and engagement in the reading. Surprisingly, the readers in this study emphasized bibliographic information over the content in each task where a genuine reading interest was the point of departure for book selection. Only when the topic was pre-decided, the content description was emphasized over title and other bibliographic information. This suggests that selecting topical novels resembles searching for non-fiction books and offers an interesting perspective to the discussion (e.g. Macgregor and McCulloch, 2006; Spiteri, 2009) on the role of professional metadata and user-generated content in library catalogs. Confirming the notions of Goodall (1989) and Saarinen and Vakkari (2013), it seems that in library catalogs, a combination of a user's previous knowledge and bibliographic information such as a title, trigger the preliminary interest and a curiosity to examine a novel in detail, whereas the content description, provided either by library professionals or other users, determines the final decision for selecting a book. The richer the metadata on item pages, the easier the decision for selecting a particular item is to make. The outcome of incomplete or missing metadata on an item page is commonly a book rejection, as the user fails in interpreting the possible reading experience. This is consistent with the findings of McKay *et al.* (2012).

Book selection as a context-related process
The fiction book selections in library catalogs were found to be highly context related. According to the search task, different expectations toward the novels to be selected were created and the novels found and encountered were assessed in the light of these expectations. Readers applied the interest criteria in a flexible way: the book selection process turned out to be a combination of affective, personal, sociocultural and metadata-related factors. The results suggested that when the author was given but previously unread, the first phase of the selection was to browse for an appealing title. The second phase was to search for novels published recently followed by examining the content description to identify an interesting item. If the author's literary production was known, the first stage was to select the most recently published title or well-known favorites to be re-read. After that, the next stage was to select engaging titles followed by discovering interesting themes from the content description.

In selecting topical novels, an important starting point for the search was participants' prior literary knowledge which was used in designing suitable queries and in identifying suitable titles from the search results. Topical search differed from the other search tasks so that together with familiarity, a dominant interest criterion was the content description instead of bibliographic information and engagement in the reading. Contrary to Koolen *et al.* (2012), novels' topical relevance was detected from

professional metadata in content description. Emphasis on the content of the novels and the irrelevance of liking in participants' interest criteria suggest that the selection of topical novels resembles that of non-fiction documents.

Differentiating from selecting known items and topical novels, selecting interesting novels in explorative browsing tasks proved to be a complex and a multi-phased process during which personal, affective and situational factors functioned as the most influential factors; serendipity was continuously involved as the readers rarely knew what they wanted to find; and the satisfying result would have been impossible to predict at the beginning of the search task. Similar phenomenon has been noted by Koolen *et al.* (2015) who found that book requests for social book search were often a combination of content, context and examples. Overall, the results yielded that explorative browsing for novels was perceived as a challenging search task because it did not provide a particular, clearly defined starting point for the search. Even though difficulty was often experienced at the beginning of the search task, participants aimed at selecting truly engaging titles by applying their literary capacity in the book selection process.

Selecting fiction books with similar reading experience in library catalogs was found to rely greatly on novels covering similar topics. Readers often settled for a topical similarity yet a mood or a particular reading experience might have been the primary criterion for the similitude. Sometimes, recommender systems fail to create recommendations based on the affective side of books. Thus, novels providing similar reading experiences to a previously read one are challenging to detect in library catalogs.

Finally, the findings yielded that compared with book selection, rejection was a quick and a straightforward process where a single attribute in bibliographic information or content description determined the novel to be non-interesting. The findings in Vakkari *et al.* (2014) and Mikkonen and Vakkari (2015) are in line with the previous as assessing novels as non-interesting was found to require less time compared with assessing them as somewhat or very interesting.

Enriched catalog supporting engagement in the book selection
The study observed the use of a traditional and an enriched online library catalog. The results suggest that interest criteria in selecting novels varied by catalog. These differences concerned familiarity, bibliographic information, content description and engagement in reading. In general, the readers in Sata emphasized the association of prior knowledge and bibliographic information in selecting novels, whereas the readers in Sampo highlighted the influence of content description and engagement in reading in their book selections. As a traditional online library catalog, Sata offered little support for readers to select novels other than known items. The users of Sata were to rely on their own ideas and literary competence for good reading as the catalog did not provide any book recommendations, lists of novel publications or diverse navigation features. As commented on by the participants, topical keyword searchers were perceived as an unnatural search tactic and metadata was often limited and displayed in an unappealing way. Lists of subject terms were often the single attribute providing information for creating the image of the expected reading experience. This explains why familiarity and bibliographic information (the title) were the most influential factors in readers' book selections in Sata.

In Sampo, the enriched features at the starting page, in search results and at item pages distinguished the book selection process greatly from Sata. Sampo offered a

continuous stream of possibly engaging reading via visual and social navigational tools, which may have altered the search process as more serendipitous compared to relying on known authors and titles. This may reduce the role of the familiarity in book selections. The enriched item pages with cover images, blurbs and automatic recommendations might have provided more complementary information to readers for generating more associations of the novels compared with Sata. As a consequence, a general interest toward the novels selected might have been easier to detect in Sampo than in Sata.

Empirical research on fiction book selection in different library catalogs is scarce. The results by Vakkari and Pöntinen (2015) on the association of an enriched results list to successful book discoveries support the findings of our study. Overall, it seems that implementing diverse features to display the content of fiction books in library catalogs is associated to the book selection process in a positive way. The possible connection between the content description and the book selection process is interesting. If the enriched content description engages the user in the book search and selecting process more compared to the query-based book search in a traditional catalog, the enriched features could have an impact on enhancing the user experience in fiction book search in online library catalogs.

Limitations and implications

A few limitations should be noted. First, the sample was biased toward females. Previous studies (i.e. Ross, 2001) have shown gender to be associated to different reading interests of pleasure readers. For example, women are more likely to read a greater amount and variety of fiction books than men (Ross, 2001). In our study, women's possible wider literary knowledge might have overemphasized the importance of participants' previous literary knowledge in the results. Second, participants' greater engagement in the book selection process in Sampo might have been a consequence of different reading preferences between the two groups. Even though the pre-questionnaire did not yield significant differences in reading activity or reading preferences between the two groups, it is possible that the users of Sampo were more interested in fiction reading than the users of Sata, which might have affected the results. Third, as the systems used in the experiment were real, it was impossible to control participants' familiarity with particular books encountered during the experiment. This might have influenced the individual results. Finally, during the data collection, neither of the catalogs used in the experiment allowed users to add tags or book reviews to the book pages. The limited amount of user-generated content in the catalogs can be considered as a limitation in the identified interest criteria, as the influence on other users' tags and reviews of the book selection process cannot be evaluated. It is an essential issue in the light of recent system design, and should be of interest in further studies examining fiction book search in library catalogs.

A few suggestions on system design can be made based on the findings. The notable role of bibliographic metadata in the book selection process confirmed the importance of traditional bibliographic indexing in library catalogs. Our results suggest that in fiction searching, bibliographic information such as the number of pages or publication year partly substitute the touching of books and it is essential in selecting between interesting and non-interesting items. A book as a whole entity continues to be vital in the selection process also in library catalogs and fresh ways of displaying bibliographic information could be emphasized in fiction indexing.

When browsing the visual and social features in the enriched catalog, the participants perceived it as challenging to pick items for a detailed examination

without a specific scope to the books displayed. Thus, offering possibilities to narrow the amount of browsable items by publication year, a genre or media recommendation could enhance the browsing experience in library catalogs. On the front page, providing categories such as "This week's most viewed novels," "Novels recently in the media," "The new chick-lit novels of 2015" or "Famous books/awarded books from the past five years" could better simulate the browsing experience in a physical library.

As in Koolen *et al.* (2015), the results of this study revealed a great challenge in designing well-functioning recommender systems for fiction books. As the previous reading experiences were often the most influential factor in readers' book selections, a recommender system capable of taking the previous experiences into consideration would be of great support in fiction retrieval. When recommending novels within the same genre, from the same author or novels covering similar topics, the automatic recommendations seem to function mostly well. However, the shortcoming is in recommending novels with a similar mood or a narrative style or creating recommendations based on a combination of various similarity criteria. The designing of a recommender system capable of combining unknown wishes, emotions and personal expectations of a fiction reader in book selection is a challenging issue in future studies.

Conclusions

The aim of this paper was to investigate fiction readers' interest criteria for novels in library catalogs. A categorization of readers' interest criteria was created based on conversations and interviews with 80 genuine fiction readers. The study shows the basic characteristics of fiction book selection in various search tasks in library catalogs. The study confirmed that the fiction readers' book selection process in library catalogs resembled that in physical libraries. Readers' prior literary knowledge and the novel's title had a major role in book selection. However, book selection in library catalogs was found to be a highly context-related process, where a combination of readers' search capacities, "behind the eyes" knowledge, personal and affective factors and a well-functioning interaction with the system being used resulted in a successful book selection. Diverging from the previous, the rejection of a novel turned out to be a quick and simple process where a single (or missing) metadata element determined a book to be rejected.

The study revealed that selection criteria varied by search tasks. Readers applied the identified interest criteria in a flexible and multiphase way depending on the search task and phase. It was found also that differences in metadata elements and content description between the catalogs were associated to the book selection process. Depending on the catalog used, the emphasis on selecting interesting titles was on either searchers' literary knowledge and bibliographic information, or content description and engagement in reading. The possible connection between the rich content description and book selection process suggested that users' engagement in selecting books in library catalogs could be enhanced by implementing rich metadata into search results and item pages.

A profound understanding of readers' book selection behavior in library catalogs is vital as fiction is increasingly accessed in digital environments. Designing well-functioning and user-friendly interfaces for fiction readers with varied search skills and literary interests requires both in-depth qualitative and large scale evaluative user studies. The findings of this study are an important step toward a comprehensive understanding of readers' book selection behaviors in digital environments.

References

Adkins, D. and Bossaller, J.E. (2007), "Fiction access points across computer-mediated book information sources: a comparison of online bookstores, reader advisory databases, and public library catalogs", *Library and Information Science Research*, Vol. 29 No. 3, pp. 354-368.

Agosti, M., Fuhr, N., Toms, E. and Vakkari, P. (Eds) (2014), "Evaluation methodologies in information retrieval", Report from Dagstuhl seminar 13441, Dagstuhl Reports Vol. 3 No. 10, available at: http://drops.dagstuhl.de/opus/volltexte/2014/4433/ (accessed October 8, 2015).

Borlund, P. (2003), "The concept of relevance in IR", *Journal of the American Society for Information Science and Technology*, Vol. 54 No. 10, pp. 913-925.

Buchanan, G., McKay, D. and Levitt, J. (2015), "Where my books go: choice and place in digital reading", *Proceedings of Joint Conference on Digital Libraries 2015 in Knoxville, Tennessee, USA, ACM, New York, NY*, pp. 17-26.

Cooper, W.S. (1971), "A definition of relevance for information retrieval", *Information Storage and Retrieval*, Vol. 7 No. 1, pp. 19-37.

Eden, B. (2002), "Using metadata to build an enriched library catalog", *Library Technology Reports*, Vol. 38 No. 5, pp. 65-70.

Goodall, D. (1989), "Browsing in the public libraries", LISU Occasional paper No. 1, Library and Information Statistics Unit, Loughborough.

Hayles, K.N. (2008), *Electronic Literature: New Horizons for the Literary*, University of Notre Dame, Notre Dame, IN.

Hypén, K. and Mäkelä, E. (2011), "An ideal model for an information system for fiction and its application: Kirjasampo and semantic web", *Library Review*, Vol. 60 No. 4, pp. 279-292.

Koolen, M., Kamps, J. and Kazai, G. (2012), "Social book search: comparing topical relevance judgements and book suggestions for evaluation", *Proceedings of CIKM'12, ACM, New York, NY*, pp. 185-194.

Koolen, M., Bogers, T., van den Bosch, A. and Kamps, J. (2015), "Looking for books in social media: an analysis of complex search requests", *Proceedings of ECIR'15, LNCS 9022, Springer*, pp. 184-196.

McKay, D., Buchanan, G., Vanderschantz, N., Timpany, C., Cunningham, S.J. and Hinze, A. (2012), "Judging a book by its cover: interface elements that affect reader selection of ebooks", *Proceedings of OzCHI'12, ACM, New York, NY*, pp. 381-390.

Macgregor, G. and McCulloch, E. (2006), "Collaborative tagging as a knowledge organization and resource discovery tool", *Library Review*, Vol. 55 No. 5, pp. 291-300.

Miall, D.S. and Dobson, T. (2006), "Reading hypertext and the experience of literature", *Journal of Digital Information*, Vol. 2 No. 1.

Mikkonen, A. and Vakkari, P. (2012), "Readers' search strategies for accessing books in public libraries", *Proceedings of Information Interaction in Context 2012, ACM, New York, NY*, pp. 214-223.

Mikkonen, A. and Vakkari, P. (2015), "Books' interest grading and fiction readers' search actions during query reformulation intervals", *Proceedings of Joint Conference on Digital Libraries 2015 in Knoxville, Tennessee, ACM, New York, NY*, pp. 27-36.

Mikkonen, A. and Vakkari, P. (2016), "Finding fiction: search moves and success in two online catalogs", *Library and Information Science Research*, Vol. 38 No. 1, pp. 60-68.

Oksanen, S. and Vakkari, P. (2012), "Emphasis on examining results in fiction searches contributes to finding good novels", *Proceedings of Joint Conference on Digital Libraries 2012, ACM, New York, NY*, pp. 199-202.

Ooi, K. and Liew, C.L. (2011), "Selecting fiction as part of everyday life information seeking", *Journal of Documentation*, Vol. 67 No. 5, pp. 748-772.

Pejtersen, A.M. (1989), "The Bookhouse: modelling user's needs and search strategies as a basis for system design", Risø Report No. M-2794, Roskilde Risø National Laboratory, Roskilde.

Peters, I. (2011), "Folksonomies, social tagging and information retrieval", in Foster, A. and Rafferty, P. (Eds), *Innovations in Information Retrieval. Perspectives for Theory and Practice*, Facet Publishing, London, pp. 85-116.

Pöntinen, J. and Vakkari, P. (2013), "Selecting fiction in library catalogs: a gaze tracking study", *Proceedings of TPDL'13. LNCS 8092, Springer*, pp. 72-83.

Reuter, K. (2007), "Assessing aesthetic relevance: children's book selection in a digital library", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 12, pp. 1745-1763.

Ross, C.S. (2001), "Making choices: what readers say about choosing books to read for pleasure", *The Acquisition Librarian*, Vol. 13 No. 25, pp. 5-21.

Saarinen, K. and Vakkari, P. (2013), "A sign of a good book: readers' means of accessing fiction in the public library", *Journal of Documentation*, Vol. 69 No. 5, pp. 736-754.

Saracevic, T. (1996), "Relevance reconsidered '96'", in Ingwersen, P. and Pors, N.O. (Eds), *Proceedings of CoLIS2, Second International Conference on Conceptions of Library and Information Science: Integration in Perspective in Copenhagen, Royal School of Librarianship, Copenhagen*, pp. 201-218.

Spiller, D. (1980), "The provision of fiction for public libraries", *Journal of Librarianship and Information Science*, Vol. 12 No. 4, pp. 238-266.

Spiteri, L.F. (2009), "The impact of social cataloguing sites on the construction of bibliographic records in the public library catalogue", *Cataloging & Classification Quarterly*, Vol. 47 No. 1, pp. 52-73.

The Reading Agency (2013), "Library facts", available at: http://readingagency.org.uk/news/library-facts004/ (accessed October 10, 2015).

The State of America's Libraries (2014), "A report from the American Library Association", available at: www.ala.org/news/sites/ala.org.news/files/content/2014-State-of-Americas-Libraries-Report.pdf (accessed October 10, 2015).

Thudt, A., Hinrichs, U. and Carpendale, S. (2012), "The Bohemian Bookshelf: supporting serendipitous book discoveries through information visualization", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, NY*, pp. 1461-1470.

Vakkari, P. and Pöntinen, J. (2015), "Result list actions in fiction search: an eye-tracking study", *Proceedings of Joint Conference on Digital Libraries 2015 in Knoxville, Tennessee*, ACM, New York, NY, pp. 7-16.

Vakkari, P., Luoma, A. and Pöntinen, J. (2014), "Book's interest grading and dwell time in metadata in selecting fiction", *Proceedings of Information Interaction in Context 2014, ACM, New York, NY*, pp. 28-37.

Yu, L. and O'Brien, A. (1996), "Domain of adult librarianship", in Godden, I. (Ed.), *Advances in Librarianship*, Academic Press, San Diego, pp. 151-190.

(The Appendix follows overleaf.)

Appendix

Dimension	Theme	Definition	Example utterances
Familiarity	Known author or title	Mention of a specific known title or an author, based on a personal reading experience	"Åsa Larsson, I know this author very well. The newest book is called The Secondly Deadly Sin. I haven't read that one yet" (SKe79) "The Captain's Daughter is one that I have read" (KSp40) "Of course I know this author, but I have never read anything from him" (KSk35)
	Known item from media and book awards	Mention of a specific known title or an author, based on a recommendation or a book review in media (newspaper, blog, television show, etc.) or literary awards	"I attended the Turku International Book Fair and there was a discussion on this book there. I will choose it" (KSa1) "Of course, I want to read Popula by Pirjo Hassinen which is nominated for a Finlandia Book Prize" (SKm64)
Bibliographic information	Author	Mention of the author of the book not known in advance	"Annikki Karikasniemi, I have never heard of her"(KSv46)
	Title	Mention of the title of the book not known in advance	"I choose Becoming Father and Daughter, solely based on the title" (SKk85)
	Cover	Mention of the book's cover	"The cover image affected my selection, in the cover there was an image of an old poster" (KSl36)
	Format	Mention of the format of the book	"It should be a paperback as I will put it in my suitcase" (KSk11)
	Publication date	Mention of the age of the book	"I prefer books published recently. I think I would choose a book published in the 21st century" (SKc53)
	Length	Mention of how long the book is	" 391 pages, it determines that I will not choose this one, it is way too long for a holiday trip" (KSy48)
Content	Genre	Mention of the genre of the book	"I want horror stories because they are easy and entertaining" (SKa51)
	Plot and Topic	Mention of the plot, topic or events of the story, including interpretation, expectations or questions	"I guess this is about different cultures, there are topics such as joy and shame" (SKe79) "I will take a look at that book called Faithful. This is about a life in an Orthodox monastery. I am a Christian but I am interested in other religions" (KSf6)
	Subject headings	Mention of the subject headings contained in the "Subject headings" field	"I will choose this because there are colonialism and cultural differences in the subject headings" (KSf6)
Engagement	Interest/Liking	Mention of general interest, like or dislike in the book	"This is absolutely brilliant, fascinating, an incredible topic" (KSp40)

Table AI.
Coding scheme for dimensions

Dimension	Theme	Definition	Example utterances
			"I wonder would it be boring to read Baby Jane, as I am not particularly really interested, but a little interested anyway […]" (KSl36)
	Mood	Mention of the mood of a particular book and mention of a wished mood for a book	"Definitely this one, this seems to be totally senseless" (KSq17)
			"I want to be entertained, I don't want to become anguished" (KSq17)
	Serendipity	Mention of an unfamiliar and serendipitous discovery	"It says Mika Waltari of France. Ooh, I can't believe this! Is this possible? What an amazing discovery!" (KSe5)
			"These three were found by chance, and actually I found more interesting titles that I expected" (SKm87)
Sociocultural	Social book sharing	Mention of friends, family members, teachers, etc. who have recommended the book not known in advance	"My mother has this book at home, I have never read it. She has recommended this to me so I will choose it" (SKv96)
	Personal connection	Mention of personal connection to some aspect of the book	"This says that the milieu of the book is Tampere and it is about students and universities. I am interested as I have studied and lived in Tampere for a year now" (KSc3)
	Intertextuality and similarity between various media formats	Mention of another book, television show or movie	"This is 14 knots to Greenwich. The topics are universities, Great Britain, stepbrothers. It reminds me of the book called Brideshead Revisited, there could be similarities" (SKn88)
			"I will choose Berlin Poplars by Anne B. Ragde. I have watched the television show based on this novel" (KSl36)

Table AI.

About the authors

Anna Mikkonen is a PhD Student at the School of Information Sciences, University of Tampere, Finland. Her research interests are on user behavior in fiction book search in digital libraries, and she has published in Proceedings of Information Interaction in Context 2012 and 2014 and in Proceedings of Joint Conference in Digital Libraries 2015. Anna Mikkonen is the corresponding author and can be contacted at: anna.mikkonen@uta.fi

Pertti Vakkari is a Professor Emeritus of Information Studies at the School of Information Sciences, University of Tampere, Finland. His research interests include fiction searching, information seeking, the growth of theories in information studies, subject knowledge and query expansion based on relevance feedback, task-based information searching and use of digital libraries. His work has appeared in the proceedings of numerous conferences, in edited monographs and in journals such as *Information Processing and Management*, *Information Research*, *Journal of Documentation* and *Journal of the American Society for Information Science and Technology*.