Hanna Jochmann-Mannak

ZOEK

Websites for children:

Search strategies and interface design

Three studies on children's search performance and evaluation

Websites for children:
Search strategies and interface design
Three studies on children's search performance and evaluation

Hanna Jochmann-Mannak

Ph.D. Dissertation Committee

Chairman and Secretary	Prof.dr.ir. A.J. Mouthaan	University of Twente, NL
Promotores	Prof. dr. T.W.C. Huibers	University of Twente, NL
	Prof. dr. L.R. Lentz	Utrecht University, NL
Assistant promotor	Prof. dr. T.J.M. Sanders	Utrecht University, NL
Members	Prof. dr. J.C. Read	University of Central Lancashire, UK
	Prof. dr. F.J.M. Huysmans	University of Amsterdam, NL
	Prof. dr. M.D.T. de Jong	University of Twente, NL
	Prof. dr. V. Evers	University of Twente, NL
	Prof. dr. F.M.G. de Jong	University of Twente, NL
		Erasmus University Rotterdam, NL

UNIVERSITY OF TWENTE.

Universiteit Utrecht

The research reported in this dissertation has been carried out at the Human Media Interaction group of the University of Twente and at the Language and Communication group of Utrecht University

CTIT

CTIT Ph.D. Thesis Series No. 13-289
Centre for Telematics and Information Technology
P.O. Box 217, 7500 AE
Enschede, The Netherlands
ISSN: 1381-3617

SIKS

SIKS Dissertation Series No. 2014-04
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

sectorinstituut openbare bibliotheken

The author gratefully acknowledges the support of The Netherlands Institute for Public Libraries

PuppyIR

The author gratefully acknowledges the support of the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 231507, called 'PuppyIR'.

Printed and bound in The Netherlands by: Optima Grafische Communicatie

ISBN: 978-90-365-3602-8
DOI: 10.3990/1.9789036536028

WEBSITES FOR CHILDREN:
SEARCH STRATEGIES AND INTERFACE DESIGN

Three studies on children's search performance and evaluation

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op donderdag 9 januari 2014 om 14.45 uur

door

HANNA EWOUDIA JOCHMANN-MANNAK

geboren op 6 januari 1983

te Rotterdam

This thesis has been approved by:

Prof. dr. T.W.C. Huibers, University of Twente (promotor)
Prof. dr. L.R. Lentz, Utrecht University (promotor)
Prof. dr. T.J.M. Sanders, Utrecht University (assistant promotor)

Contents

ZOEK

Chapter 1

Introduction

The mother of an eight year old told us: "My daughter got an assignment at school to give a presentation for her classmates. She decided to give a presentation about rabbits. So, she asked me to help her find information about rabbits on the internet. We took the laptop and I asked her: "What do you think we should type into Google to find information about rabbits?" and she said: "Well, I would type: In-for-ma-tion"!

This is one of the many anecdotes - and for us the most striking one - that people shared with us when we told them about the type of research that we were working on: children's digital information search. This anecdote exemplifies the necessity to study children's information search and to learn more about how children think and how digital search environments should be designed for children to support them in effective information search. Our research is about children in the age of 8 to 12 years old who are able to read and write.

1.1 Children's information search behaviour

Every day, digital media play a more important role in children's lives. Besides playing games, children use digital media for social gatherings, to watch videos, to make their own digital space and as a support for learning. The number of technological services on the internet and apps for iPads or other tablet computers, especially designed for children, is growing exponentially. And although we tend to assume that children know how to work with computers and digital interfaces – they were practically born with them – the reality of children's search performance is less positive.

The question we were most frequently asked while we worked on this research project was: "Why is this necessary? Isn't Google the answer to all of your questions?" And indeed, up to 80% of the children in the Netherlands already use Google or Google-type search interfaces to find information on the Internet from an early age. However, the fact that children use Google does not automatically mean that search interfaces such as Google are most suited for children's information needs and skills. The most important problem for children with search interfaces such as Google is that these interfaces are not designed for children. Both interaction design and query matching algorithms are developed for adults. Therefore, information might not be relevant or may not even be safe for children. Children have problems with deciding what is relevant and which sources are reliable or not. Search interfaces need to be designed for or adapted to children to provide a safe environment especially for children. Therefore, the question arises how should these search interfaces be designed or adapted to children. Is it indeed the case that search interfaces such as Google are not good enough for children?

Prior research identifies all kinds of problems children encounter during information search, when they are confronted with these search interfaces that are designed by and for adults (Druin, Hutchinson, Foss, Hatley, Golub, Leigh Guha & Fails, 2009). For example, children have trouble formulating search queries or selecting relevant subject categories, because they have less domain knowledge than adults and a less developed vocabulary. Another example of a problem is that children have difficulties judging the relevance of search results. To support children in effective information search, search systems need to be designed that are consistent with children's needs and cognitive skills. That is why research needs to be conducted to find out how children's needs and skills can be supported by digital interfaces.

The aim of the research reported in this dissertation is to learn more about children's information search behaviour and search interface interaction and to find out how to design interfaces for children that 'work' for children (i.e. effective information search) and how to design interfaces that children 'like' (i.e. positive evaluation). This takes us directly to the first two objectives of this dissertation. First of all, the research concerned children's information search behaviour, which means that it was about child-interface interaction in an informational and educational context. The social and entertainment contexts are outside the scope of this dissertation. Secondly, the research was about *children's* information search behaviour, which means that it was about children that can read and write, but more importantly, about children that mostly comprehend what they read and write. Therefore, our research focus was on children between 8 and 12 years that are able to read and write. Children that experience 'technical' problems with reading and writing were outside the scope of our research.

Many synonyms are used for the 'object' of our research: interfaces, systems, digital environments or websites. To remain consistent throughout this dissertation, we will continue using the term 'interface' or 'search interface' when talking about the interfaces of the search systems used by the children in the studies presented in this dissertation.

The idea for this research project was conceived by the Dutch Library Association in 2008 in their attempt to design a digital interface for children, called the 'Digital Youth Library'. They soon realized that there was not much knowledge about children's digital information search behaviour and that there were no clear guidelines for the design of children's digital interfaces. They saw the necessity to set up a research project alongside this design project so that their work could be supported by more fundamental research.

1.2 Relevant research traditions concerning children's information search

This research project was multi-disciplinary, because theories and results of different research traditions are relevant and indispensable for a thorough study of children's search behaviour. The most important research tradition for this dissertation is the human-computer interaction community and more specifically, the child-computer interaction or interaction design and children community (IDC). In this community, research is conducted on children's interaction with all types of technology. Most research by this community is aimed at novel and new emerging technologies for children (such as multi-touch tables (Rick, Harris, Marshall, Fleck, Yuill & Rogers, 2009), or tangible interfaces (Xu, Read, Mazzone, MacFarlane & Brown, 2007; Detken, Martinez & Schrader, 2009). Other research is aimed at technology that belongs to the desktop paradigm (Druin et al., 2009) to which this dissertation belongs. The IDC-community is relevant for this dissertation, because it provides models that help designers and researchers to better understand children and their relationship with technology and methods for working with children in the design and evaluation studies of technologies designed for children.

More specifically than child-computer interaction, this research was about children's information search (which includes both searching and browsing) within digital environments. Therefore, research traditions that concern models for information search are relevant for this dissertation. A research tradition that works with information seeking models from a user point of view, is the Library and Information Science research tradition (Kuhlthau, 2004; Dresang, 2005). A research tradition that works with information seeking models from a system point of view, is the Information Retrieval research tradition (Van Rijsbergen, 1979; Croft, 1993; Manning, 2008). A research tradition that integrates this user and system point of view is the integrated information seeking and retrieval research tradition (Ingwersen & Järvelin, 2005; Lazonder & Rouet, 2008). We considered that the models from this research tradition were most suited for our research on children's information seeking, because all relevant factors that influence the information seeking process are involved in this research tradition. This will be discussed in more detail in Chapter 2.

We studied children's interaction with digital interfaces from two perspectives:

1) from a *performance* perspective, which is about pragmatic factors, such as search effectiveness and search efficiency (Borgman, Hirsh, Walter & Gallagher, 1995; Schacter, Chung & Dorr, 1998; Bilal, 2000; Druin, 2005)
2) from an *evaluation* perspective, which is about children's User Experience (Hassenzahl, 2004; Thüring & Mahlke, 2007; Van Schaik & Ling, 2008; Tuch, Roth, Hornbaek, Opwis & Bargas-Avila, 2012), such as their perception of instrumental

and non-instrumental qualities, aesthetics, emotional reactions and overall appraisal of a search interface.

Children's evaluation of a search system might not only be based on pragmatic issues of a search interface, but also on the children's feelings towards an interface. This will be discussed in more detail in Chapter 2.

1.3 Main research questions in this dissertation

The main objective of this research was to learn more about children's information search behaviour and how search interfaces can support children in effective information search. The following research questions (*RQs*) are the main questions answered in this dissertation.

To study children's information search behaviour, we wanted to know whether there were specific design principles for children applied in search interfaces for children, anno 2005-2010.

RQ 1. *Which design conventions on interfaces for children are specific for children and which of them are general design conventions?*

When we know whether there are specific design conventions for children and what they are, the next question is which design characteristics do and which design characteristics do not work for children's search performance. Another question is whether search interfaces that work well for children are also appreciated by children. In other words, which design characteristics of search interfaces have a positive effect on children's attitudes towards and evaluation of these search interfaces? Is children's performance on a search interface related to their affective responses towards a search interface?

RQ 2. *What is the effect of design characteristics of informational websites for children on their search performance and on their appreciation of the search interface?*

RQ 3. *What is the relation between children's search performance and their affective responses towards the search interface?*

Finally, a common issue on children's search performance is the search strategy that is suited for children to search for relevant information. The two most frequently used and studied search strategies are keyword searching (such as the Google-type searching) and browsing subject categories. Therefore, we also studied this issue in our research.

RQ 4. *What is the difference between browsing and searching in terms of efficiency and effectiveness of children's search task performance?*

1.4 Mixed method research design

In principle, within the Humanities and Social Sciences, we can distinguish between three types of research designs: qualitative, quantitative and mixed methods. Mixed methods are also sometimes referred to as multi-method approach, multimodality or triangulation (Nisbeth Jensen, 2013). A comprehensive definition of mixed methods is provided by Creswell and Plano Clark (2007):

"A research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis and the mixture of qualitative and quantitative approaches in many phases of the research process. As a method, it focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches, in combination, provides a better understanding of research problems than either approach alone." (p. 5)

To answer our research questions, we did indeed opt for a mixed methods design, following Creswell and Plano Clark (2007), because both quantitative and qualitative methods were needed as the research design involved both exploratory and empirical elements as will be described below. More specifically, this research followed a pragmatic approach using a sequential mixed methods design with three studies, and therefore, the research questions continually evolved and became more specific based on the findings in each of the three studies as these three studies built on each other.

"Iterative sequential research designs are characterized by the mixing of qualitative and quantitative methods which occurs in a dynamic, changing, or evolving manner over the course of the research project, meaning that the findings at one stage influence decisions about the next stage." (Nastasi, Hitchcock & Brown, 2010, p. 320)

In short, we would like to present our mixed methods design as a 'triple jump'-design; three iterative steps following one another in time to reach our research goals. Figure 1.1 shows an overview of our overall research design, which will be described in more detail below.

Step 1 (hop): Corpus study

Unlike other web genres (Peck & Hane, 1998; Adkisson, 2002; Herring, Scheidt, Bonus & Wright, 2004), to our knowledge the genre of children's informational websites has not yet been studied and described systematically by means of a corpus study on design characteristics. To study children's search behaviour on informational websites, it is important to learn more about the general design conventions for this web genre first. Therefore, in the first step ('hop') of our research, we conducted a corpus study with a corpus of 100 international, informational websites for children to analyze the design

Step 1 (*hop*):	Step 2 (*step*):	Step 3 (*jump*):
Corpus study to identify design conventions of children's search interfaces	**Explorative study** to explore children's search behaviour on existing search interfaces, that are representative for children's search interfaces based on the results of the corpus study.	**Experiment** to test children's search performance on and affective responses towards three types of search interfaces for children. Research directions suggested by the results of the explorative study will be empirically tested in the experiment, with interfaces that we will specifically develop for this experiment, based on the results of the corpus study.

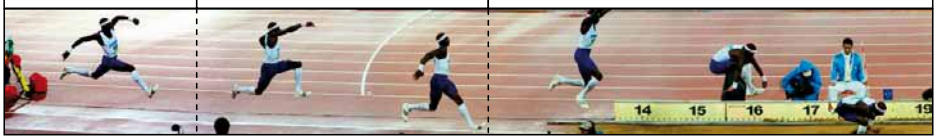


Figure 1.1: Overall sequential research design

characteristics of these interfaces and to identify which design characteristics are conventional in the genre of children's informational websites. Design characteristics were analyzed are, for example, visual design characteristics and characteristics of navigation and information architecture. Every single interface from the corpus was analyzed by two evaluators so that inter-coder reliability could be computed. In total, five evaluators (i.e. the author of this dissertation and four research assistants) worked on the analysis of the interfaces from the corpus. Based on the results of the corpus study, we studied whether there are different types of designs for children's interfaces.

We used this quantitative research method in the corpus study to analyze design characteristics of existing search interfaces for children, in order to find an answer to *RQ 1*. This is a scientifically justified method to systematically analyze a large number of children's informational websites without subjective interpretations of the findings.

Step 2 (step): Explorative study

The results of the corpus study formed the basis for the next 'step' in our research, by exploring children's search behaviour and affective responses in relation to the design characteristics identified in the corpus study. The results of the corpus study guided the selection of interfaces with which children would interact in the explorative study, to find out which factors of interface design influence children's search performance and system evaluation most. We selected interfaces for the explorative study with design characteristics that represent different types of interface design, when different types could be identified in the corpus study.

In the explorative study, we used a concurrent mixed methods design (Creswell, 2009) by merging quantitative (e.g. search performance scores for amount of time and number of clicks needed to conduct search tasks) and qualitative data (e.g. analysis of used search strategies by children) in order to provide a comprehensive analysis of

the research problem, in which both forms of data were collected at the same time. More specifically, we observed children while they were conducting search tasks on several (existing) search interfaces for children (one interface at a time) by observing and recording their screen activities. We also observed and recorded children's eye movements on the screen while they searched for information. Further, we automatically logged their search activities by recording amount of time, number of clicks and other search events (e.g. typing a query) used to find relevant information. Children's evaluations of the interfaces were assessed using different evaluation methods (e.g. by using smiley scales and by 'ranking' the interfaces).

Step 3 (jump): Experiment

Based on the insights that were provided by the corpus study and the explorative study, we set up an experiment in which we empirically tested the hypothesized effects (from the explorative study) of the design characteristics (that were identified in the corpus study) on children's search performance and user experience. In this experiment, we also used a concurrent mixed methods design by merging quantitative and qualitative data in which both forms of data were collected at the same time. More specifically, just as in the explorative study, we observed children while they conducted search tasks on the search interfaces by observing and recording their screen activities. Only this time, the search interfaces were developed specifically for this experiment, which allowed us to control for effects of the system. Further, we automatically logged their search activities by recording amount of time, number of clicks and other search events (e.g. typing a query) used to find relevant information. We did not record children's eye movements in the experiment, because 'exploring' children's screen behaviour is not an objective in this experiment, as was the case in the explorative study. Children's evaluations of the interfaces were assessed using different evaluation methods, based on the experience we had gained with these methods during the explorative study. Both the results of the explorative study and the experiment enabled us to answer *RQ 2* and *RQ 3*. Further, in the experiment, we also controlled the use of the search engine versus the use of the main categories to search for information, which enabled us to answer *RQ 4*.

The fact that the studies that we conducted also enabled us to evaluate which research methods are suited best for research with children is of added value to our research. Both objective and subjective measurements to study children's search behaviour, search performance, subjective perceptions and evaluations of search interfaces and affective responses towards these interfaces have been evaluated in our research project.

1.5 Chapter overview of this dissertation

Chapter 2 in this dissertation positions the research in the existing research traditions about information search. We will provide an overview of factors that are relevant to children's information search behaviour and what the literature reports concerning these factors. We will also define the factors that we analyze, evaluate and manipulate in our research. Finally, we will present the two perspectives that we used in our studies on children's search behaviour, the *performance* perspective (i.e. pragmatic usability scores such as effectiveness and efficiency) and the *evaluation* perspective, inspired by the User Experience research tradition (Hassenzahl, 2004; Thüring & Mahlke, 2007; Van Schaik & Ling, 2008; Tuch et al., 2012). This will be followed by an overview of our own research objectives in this research project.

In Chapter 3, we will give an answer to the first research question by describing what the genre of children's informational websites looks like *(RQ 1)*. We will report a corpus study and give an overview of design conventions of children's informational websites. This study showed that design conventions used for children are not very different from design conventions used for adults. However, while Chapter 3 will broadly show no differences between design conventions for children and adults, a more detailed analysis of the results, at the end of Chapter 3, will show that informational websites for children can be categorized as three types of children's websites.

Chapter 4 will describe a first exploration into children's information search behaviour, both by browsing and by searching *(RQ 4)*. Children's search behaviour have been studied on three types of informational websites; three existing Dutch websites, designed for children *(RQ 2)*. The selection of these interfaces was based on the results of the corpus study (Chapter 3). This behaviour was also studied in relation to children's search behaviour via Google. This explorative study gave signals about the important role of the affective side of children's experience while using these websites *(RQ 2)*. Therefore, this study was a first exploration of both the *performance* perspective on usability and the *evaluation* perspective on children's user experience with digital interfaces and the relation between these 'two sides of the same coin' *(RQ 3)*. In addition, the study was also a first exploration in the type of methods that are suited for research on children's interaction with digital interfaces. The results from the research that will be reported in Chapter 4 will give suggestions about which interface design factors are of real importance for children's information search behaviour.

We conducted empirical research on the effects of some of these interface design factors, which we will report in Chapter 5. The factors were not included as we came across them in their natural environment (as is the case in Chapter 4), but were manipulated in experimental stimuli, especially designed for this empirical study. These manipulated factors were the amount of playfulness in the visual and navigation design

of the interface and the use of the search engine versus the use of the navigation structure to search information. The 'ideas' about which interface design factors are important in children's information search behaviour from Chapter 4 and the important role of the affective side of children's experience while using these websites, were tested empirically in this study. Effects of these factors were measured from both a *performance* perspective and an *evaluation* perspective of children's information search behaviour. From a performance perspective, factors were measured such as the amount of time and number of clicks needed to find relevant information or the success with which relevant information was found. From an evaluation perspective, factors were measured such as children's perceptions of the quality of the interfaces, their emotions towards the interfaces and their final evaluations of the interfaces in terms of beauty, goodness and fun *(RQ 2)*. Also the relation between pragmatic performance and affective responses was studied *(RQ 3)*. Further, both key word searching and browsing subject categories were studied in the experiment *(RQ 4)*. Finally, suitability and validity of methods to measure pragmatic performance and affective responses of children will be evaluated in Chapter 5.

In Chapter 6, we will summarize the main findings and we will use these findings to answer the main research questions of this dissertation and to discuss the role of interface design and search strategies during children's information search. We will also return to issues raised in Chapter 2 and relate our research to theories from prior research. In addition, we will reflect on the methods used in our study and the suitability of these methods in research with children. Finally, we will discuss the results of our research in terms of the future of children's informational interface design and we will provide suggestions for future research.

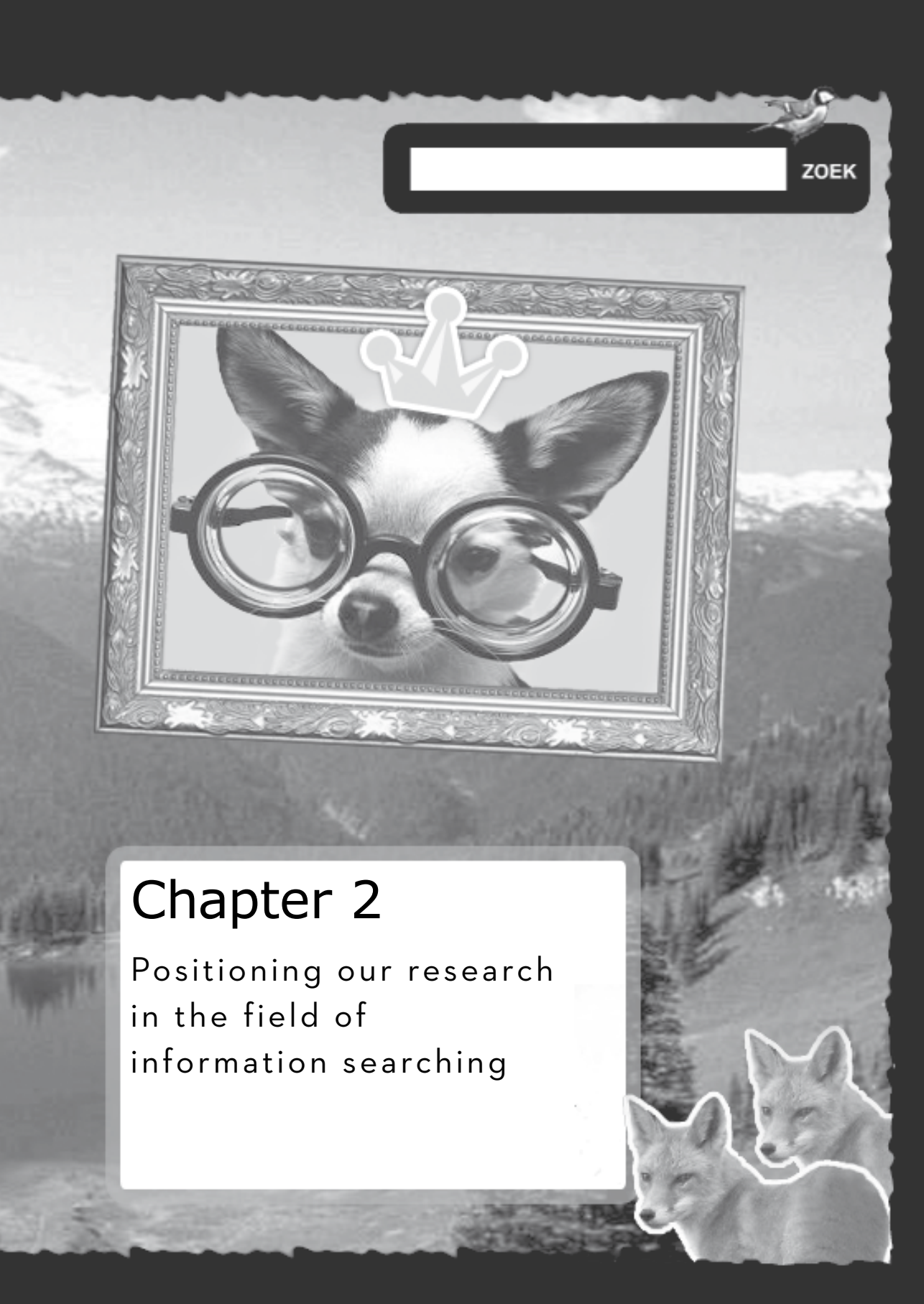
1.6 Reading guide

There is some overlap in the literature discussed in the various chapters of this dissertation. This overlap is due to the fact that the various chapters have been set up as separate journal articles. We trust this redundancy is of help to the reader, because each chapter can be read on its own.

Chapter 3 was published in *Technical Communication* (Jochmann-Mannak, Lentz, Huibers & Sanders, 2012) and honored with the Frank R. Smith Outstanding Journal Article Award (2013) of the Society for Technical Communication and the Silver Excel 2013 Award in the category 'Journal: Feature Article'. Chapter 4 is based on a paper that was presented at the Workshop on Accessible Search Systems at the 2010 ACM SIGIR conference that was held in Geneva, Switzerland. Chapter 5 is published as a

book chapter in "Evaluating websites and Web Services: Interdisciplinary Perspectives on User Satisfaction"(Yannacopoulos, et al., 2013)[1].

1. Preliminary versions of the chapters in this dissertation have been presented at several conferences. The references are listed in the Publications section at the end of this dissertation.



Chapter 2

Positioning our research in the field of information searching

In this chapter, we will position our research in existing research traditions, we will discuss factors that influence children's information search behaviour according to previous research and we will discuss variables measured in previous research and methods used to measure these variables. We will present these findings from previous research to give a full understanding of the domain of children's information search and retrieval that helped to guide our decisions for the main research objectives in our research project. The main research objectives will be summarized at the end of this chapter.

2.1 Models of information behaviour

Before describing models of various research traditions that are relevant for our research project, we will present Wilson's (1999) Nested Model of Conceptual Areas (see Figure 2.1). Analysis of various information seeking and information search models led Wilson (1999) to suggest the following:

"The various areas of research within the general field of information behaviour may be seen as a series of nested fields. Information behaviour may be defined as the more general field of investigation (as shown in Figure 2.1), with information seeking behaviour being a sub-set of the field, particularly concerned with the variety of methods people employ to discover and gain access to information resources. Information searching behaviour is then defined as a sub-set of information seeking, particularly concerned with the interactions between information user (with or without an intermediary) and computer-based information systems, of which information retrieval systems for textual data may be seen as one type." (p. 269)

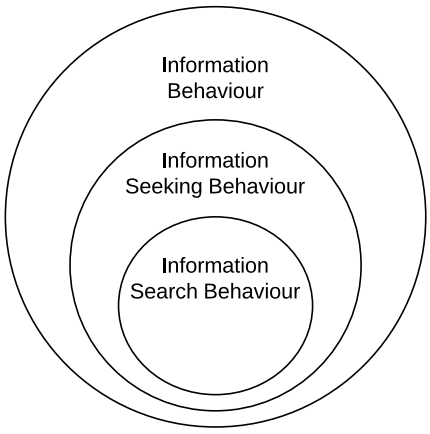


Figure 2.1: Wilson's Nested Model of Conceptual Areas (Wilson, 1999)

In this same article, he concluded:

"The various models of information behaviour, information seeking behaviour and information searching behaviour represent different aspects of the overall problem: they are complementary, rather than competing." (p. 269)

We think that Wilson's view on information behaviour is enlightening and therefore, we followed his view in our research on children's information behaviour.

The core of our research project was about children's information searching behaviour, which can be seen as a sub-set of the field of information seeking behaviour. Information searching here was particularly about children's interaction with a computer-based information system to search for relevant information during the information seeking process. However, we think it is important for our research to see the bigger picture and to take into account the larger context of children's digital search behaviour. Therefore, in line with the insights from Wilson's model, we will first discuss models of information seeking behaviour and after that more specific models of information searching behaviour (e.g. traditional Information Retrieval models).

We will start by presenting several models from prior research on children's information seeking and information retrieval to describe how our research can be seen in relation to these models. These theories about (children's) information seeking and information retrieval are derived from Library and Information Science (Kuhlthau, 2004, Cooper, 2002), Information Studies, information retrieval (Manning, 2008), Instructional Technology (Lazonder & Rouet, 2008), Information Seeking and Retrieval research (Ingwersen & Jarvelin, 2005; Tanni, 2008) and Human-Computer Interaction design (Druin, 2003).

2.1.1 Models of information seeking in Library and Information Science

Educators and librarians have designed various models to illustrate how teachers and learners act in information seeking or problem-solving situations. In the 1990s, the models began to stress the ongoing cycle of inquiry. Rather than a series of separate steps, educators began to see the process as involving recursive elements and ongoing questioning, exploration, and investigation. In all these models the process of information seeking is described in different phases. The most commonly used model is Kuhlthau's (2004) Information Seeking Process (ISP) model (see Figure 2.2). The ISP model consists of the following steps. The process starts with a shortage of information (information gap, information need) in which the user 'initiates a task' and 'selects a topic'. Then the user refines the problem by searching, Kuhlthau calls this 'formulation'. After that, the user gathers relevant information to solve the problem. This is called 'information collection'. And finally, the user can start writing or presenting the task output; the so-called 'presentation'. The stages in the search process are iterative in

Tasks	Initiation	Selection	Exploration	Formulation	Collection	Presentation
Feelings (affective)	uncertainty	optimism	confusion / frustration / doubt	clarity	sense of direction / confidence	satisfaction or disappointment
Thoughts (cognitive)	vague →			focused →	increased interest →	
Actions (physical)	seeking relevant information exploring →				seeking pertinent information documenting	

Figure 2.2 Model of the Information Search Process (ISP) (Kuhlthau, 2004, p.82)

nature, which means that different stages might be repeated within the search process. This model is not specifically related to children. However, we do think that this model is also relevant for research on children's information seeking behaviour to analyze whether children experience the same feelings, thoughts and actions as adults in the information seeking process.

Baker (2005) reported the results of a comparison between the most common information seeking models (e.g. 8 W's, Big6, ISP, I-Search, Pathways, Research Cycle and REACTS). She found that the steps of most models could be organized in the following phases. However, the authors of the models all use their own terminology for the different phases as reported per phase in parentheses.

1. Questioning (task definition, initiation, defining, formulating, choosing a topic, etc.)
2. Exploring (seeking, planning search strategies, locating, search, etc.)
3. Assimilation (information collection and use, gathering and integrating, interpretation, synthesizing, analyzing)
4. Inference (communication, representing knowledge, organize and present, transforming, reporting)
5. Reflection (evaluation of product and process, search closure).

In our research, we will see whether these phases can also be recognized in children's information search behaviour.

Strengths and limitations of information seeking models for our research

The models focus on the behaviour and characteristics of the user and the way the user can be educated or instructed to improve the information seeking (or problem-solving) process (Walraven, Brand-Gruwel & Boshuizen, 2008). The main limitation of these information seeking models for our research was that they ignore the influence of the 'work task context' in which the information seeking process takes place. Furthermore, the models ignore the important role of the modern tools (e.g. the various websites on the internet) that support users during information seeking and information use. In our opinion, the role of the search system and the user interface, and the interaction between the user and the system are neglected in these models.

2.1.2 Models from traditional Information Retrieval research

Information Retrieval (IR) is about finding information (usually in documents) of an unstructured nature (usually text) to satisfy from within large collections (usually stored on computers) an information need (Manning, 2008). Traditional research on IR focuses on how an information need from a user can be matched with information within a system (e.g. the internet) to satisfy the user's needs by retrieving relevant information from a system (i.e. type of information search behaviour). Figure 2.3 presents the dimensions of a traditional IR model. The model represents IR as two tracks of elements and processes converging on comparison, also named 'matching'. Figure 2.3 shows a user track at the left and a system track at the right.

The user track starts with a user's information need that is verbalized in a question and then transformed and formulated into a query acceptable to the system. The system track involves documents (e.g. texts, images) that are indexed in a particular way, then organized in a file of indexed documents and in this way gets ready for comparison (i.e. matching). Then matching between the two representations (e.g. query and indexed documents) occurs. Feedback is included that allows for modification of representations, but usually modifications only occur in the user track (Croft, 1993). These basic principles of information retrieval are independent of the type of user and therefore, also relevant for users that are children.

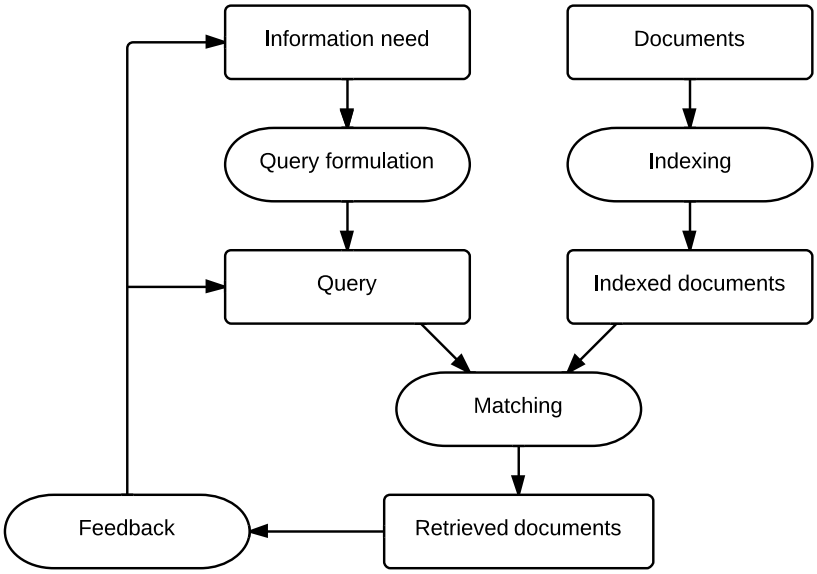


Figure 2.3: Traditional Information Retrieval processes model (Croft, 1993)

Strengths and limitations of traditional Information Retrieval models for our research

The strength of the traditional IR model is that it allows for straightforward isolation of variables and for uniform analysis of diverse IR methods and algorithms. An important limitation of this traditional IR research model for our research was that the model does not take into account the context the user is in. Ingwersen and Järvelin stated that most IR researchers develop search engines for artificial search tasks without evaluating how real people search information with the developed search algorithms or tools (Ingwersen & Järvelin, 2005). In our research, the role of the user and the context is indispensable.

2.1.3 Integrated IS&R Research Framework by Ingwersen and Järvelin

Not only factors concerning the context or user can be of influence on information searching (and in our case on children's information searching), but also factors of the system, such as the way in which information is structured by a system (e.g. through subject categories or through a keyword search tool) or the way in which information is presented visually, play an important role in children's information search. Besides the question of which steps a child should follow to effectively find relevant information and solve a problem, we should ask ourselves how a system can be improved to support a child in effective information search. In our research, we were interested in children's information interaction by using a search interface in which effects from context, task and user characteristics are just as important as effects from the system characteristics. Neither the traditional information seeking (IS) models nor the information retrieval (IR) models address all of these characteristics in one model. That is why the model of Ingwersen and Järvelin (2005) was more suited for our research (see Figure 2.5). In this model, the interaction between a child and a system during the information search process is presented, in which both characteristics of the child and characteristics of the system to support the child's search behaviour, are represented.

Ingwersen and Järvelin (2005) made an overview of the dimensions that play an important role in the domain of information seeking and retrieval. Figure 2.4 presents which of these dimensions were studied in traditional Information seeking research, and traditional (online interactive) information retrieval research. This gives a clear overview of the limitations of traditional IS&R research for research on information search behaviour (and for our research more specifically on children's information search behaviour) as described above.

Because of the limitations of traditional IS&R research for studying information seeking behaviour, Ingwersen and Järvelin (2005) made a further integration between traditional information seeking research (IS) and Information Retrieval (IR) research by presenting an integrated information seeking and retrieval (IS&R) Research Framework.

They proposed five broad categories with nine classes of variables that interact in information seeking and retrieval processes, called 'dimensions':

1. Organizational task dimensions:
   a. Work task dimension: work task (also non-job related), organization of work, collaboration and system environment.
   b. Search task: seeking and retrieval tasks
2. Actor dimensions:
   a. Actor dimension: actor's declarative knowledge and procedural skills
   b. Perceived work task dimension: actor's perception of the work task
   c. Perceived search task: actor's perception of the search task including information need types and the task performance process; emotions.
3. Document dimension: document genres and collections in various languages and media, which may contain information relevant to the task as perceived by the actor.
4. Algorithmic dimensions:
   a. Algorithmic search engine dimension: the representation of documents/ information and information needs; tools and support for query formulation; matching methods.
   b. Algorithmic interface dimension: tools for visualization and presentation.
5. Access and interaction dimension: strategies of information access, interaction between the actor and the interface.

Research Tradition / Dimension	Traditional IS Research	Traditional Online IRR Research	Traditional IR Research
Work Task Dimension	☹	⊘	⊘
Search Task Dimension	☺	😐	⊘
Action Dimension	☺	☺	⊘
Perceived Work Task Dimension	☹	⊘	⊘
Perceived Search Task Dimension	☺☺	😐	☹
Document Dimension	😐	☹	☹
Search Engine Dimension	⊘	☹	☺☺
Interface Dimension	☹	☺	😐
Access & Interaction Dimension	☺	☺☺	☹

Legend: Dimension…. ⊘ excluded from study 😐 fairly in focus of study ☹ little in focus of study ☺ strong focus of study

Figure 2.4: Foci of traditional IS&R research from Ingwersen and Järvelin (2005)

Figure 2.5 represents these dimensions of the IS&R research framework and how they relate to each other. The figure shows clearly that the information seeker is the central cognitive actor in the framework. To understand the process represented in the framework, it is useful to 'read' the framework from right to left.

I. The 'social, organizational and cultural context' on the right represents the work task dimension (IS&R dimension 1a).
II. The 'information seeker's cognitive space' in the center represents the search task dimension (1b) and the actor dimensions (2a, 2b and 2c).
III. The 'information objects' on the upper left hand side represent the document dimension (3).
IV. The 'IT' and the 'Interface' on the left hand side represent the algorithmic search engine dimension (4a) and the algorithmic interface dimension (4b) respectively.
V. Finally, the access and interaction dimension (5) is represented by the interaction arrow "R2" between the information seeker and the interface.

To conclude, Ingwersen and Järvelin (2005) provided this helpful model that was also relevant for our study on children's information searching behaviour. By studying the different dimensions of the search process mentioned in it, all possible factors that might influence children's search performance could be addressed.

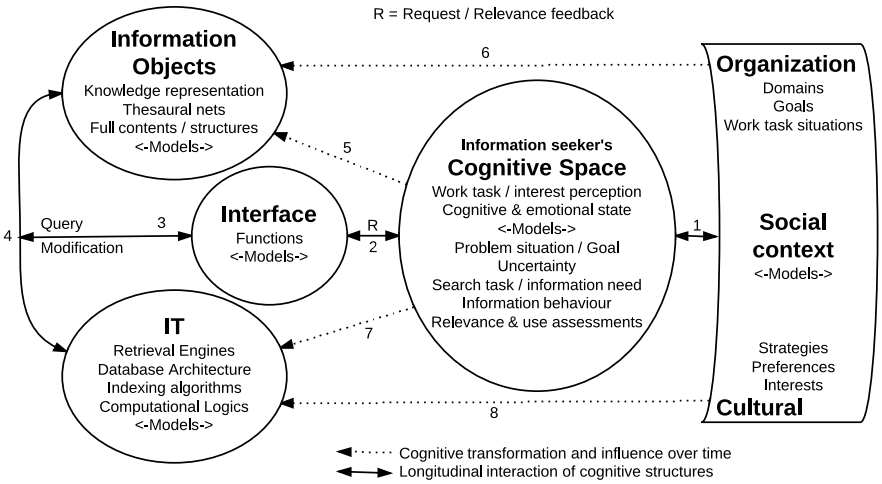


Figure 2.5: Cognitive framework of IS&R research from Ingwersen and Järvelin (2005)

2.1.4 Access and interaction dimension in more detail

In our research we were particularly interested in the actual information search process that takes places within the Access and interaction dimension (see R2 in Figure 2.5)

(Ingwersen & Järvelin, 2005). However, they did not study this 'access and interaction dimension' in more detail.

Researchers that do study the interaction between a user and a system (in particular navigating the Web) in more detail are Kitajima, Blackmon and Polson (2000). They developed a cognitive model to simulate navigation on the web, termed the Comprehension-based Linked Model of Deliberate Search (CoLiDeS). This model assumes that comprehension of texts and images is the core process underlying Web navigation and it is inspired by the text comprehension (construction-integration) theory of Kintsch (1988). More specifically, it is inspired by the concept of 'information scent' (semantic relevance of screen objects to users' goals) and emphasizes the semantic dimension of Web navigation; that is, they assume that the process of relevance assessment is central to web navigation. Information scent is measured based on three factors: semantic similarity, frequency and literal matching. Semantic similarity is calculated based on co-occurrences between words and documents with the aid of a machine learning technique called latent semantic analysis (LSA).

Juvina and Van Oostendorp (2008) show that not only semantic but also structural (spatial) knowledge is involved in navigating the Web. That is why they expanded the CoLiDeS model into CoLiDeS+ that uses 'information scent' to account for user's judgments of relevance (semantic dimension) and 'path adequacy' (the semantic similarity between a navigation path and a user's goal) to account for the user's efficiency in traversing a Web structure (structural dimension). Figure 2.6 presents the steps of the CoLiDeS+ model that an information-seeker takes during web navigation to find relevant information. A short description of these steps is presented below the figure. As can be seen in Figure 2.6, CoLiDeS is a detailed and operational model, which can be very useful for qualitative study of users' search behaviour and search strategies while they carry out search tasks within digital environments.

1. A task description is taken as input (equivalent to user's goal).
2. A web page is attended to, parsed in several areas and a particular area is focused on (e.g. a menu).
3. Menu entries are comprehended (based on how semantically similar they are to the user's goal) and the entry that is most relevant to the user's goal is clicked on.
4. A new page is attended to and if the target information cannot be found, the cycle is reinitialized (via 'back' or 're-focus').
5. The selected element is retained in a memory structure that maintains user's navigation paths. Starting with the second cycle, a navigation path is available and the metric called 'path adequacy' is computed.
6. If screen objects do not contribute to an increase in path adequacy an impasse is declared and dealt with by considering 'next best' options.

7. The algorithm stops when the user decides that the current page contains the target information.

CoLiDeS (Kitajima et al., 2000) and CoLiDeS+ (Juvina & Van Oostendorp, 2008) are not the only recently developed cognitive models aimed at a better understanding of web navigation by simulating how users navigate over a series of web pages. Another recent model is called SNIF-ACT. This is a model developed within the information foraging theory which employs the rational analysis method. To assess the utility of navigational choices, the mechanism of spreading activation is used. This means that activation from representation of information scent cues spreads to the users information goal. The amount of the activation received by the user's goal reflects the expected utility of choosing navigation actions associated with those cues (Fu & Pirolli, 2007). In modeling the role of information scent in navigation choice, SNIF-ACT is much like CoLiDeS.

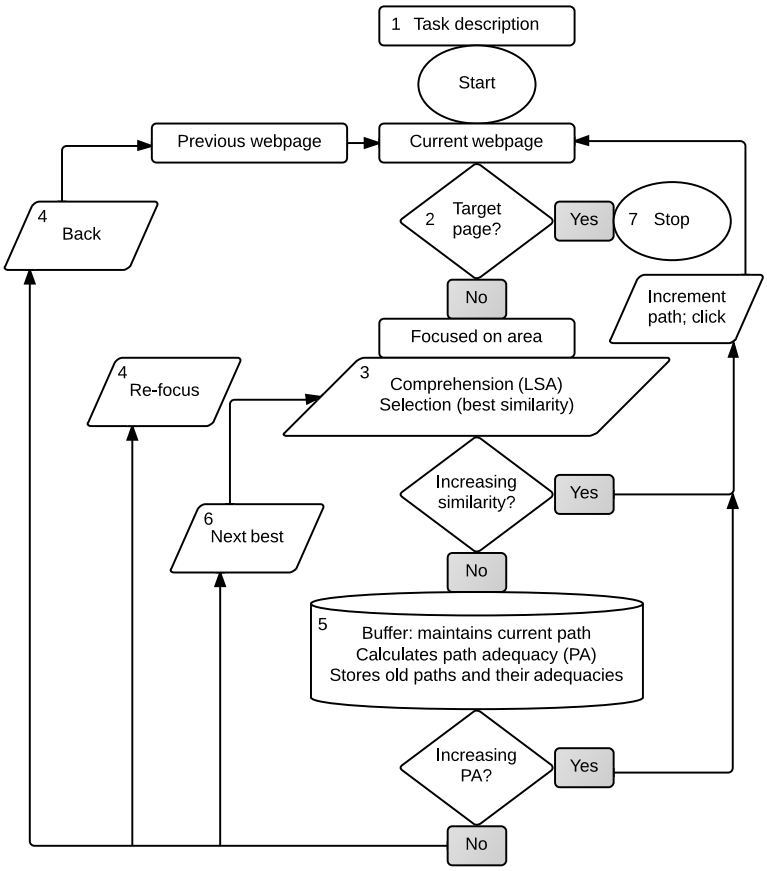


Figure 2.6: A diagram of the algorithm that implements the CoLiDeS+ model (Juvina & Van Oostendorp, 2008)

SNIF-ACT assumes that links on a web page are processed sequentially. That is realistic when considering tasks that are conducted using a search engine. However, sequential processing may not apply that well on web pages that are categorized under different headings to identify relevant subregions. This type of web pages implies a hierarchical instead of sequential processing of links. In contrast with CoLiDeS, SNIF-ACT is developed at a level of abstraction that is not sensitive to different visual layouts of web pages. CoLiDeS was developed at a more concrete level of abstraction and is therefore more suitable for a hierarchical organization of information.

2.1.5 Conclusions concerning the described models

Our first goal was to find a model which was most suited for our research to apprehend children's information interaction using a search interface. We looked for a framework for information seeking and retrieval in which both user characteristics and system characteristics are integrated. We think that both the user and the system have to be examined in research to eventually design an interface that supports the user in effective information seeking. We found such a framework in the Integrated IS&R Research Framework from Ingwersen and Järvelin (2005). However, the interaction dimension in this framework was not addressed in detail. That is why we described the cognitive models of navigating the Web, CoLiDeS (Kitajima et al., 2000) and SNIF-ACT (Fu & Pirolli, 2007). There are several reasons why the CoLiDeS-model was interesting to guide the observations of children's interactions with the interfaces in our research.

- Both human-media interaction research and discourse research is addressed with this model, because CoLiDeS is inspired by the text comprehension (construction-integration) theory of Kintsch (1988).
- The model offers opportunities for experimental research and the model is useful to observe and analyze search behaviour and search strategies.
- The model offers opportunities to vary user characteristics (e.g. cognitive characteristics such as scanning skills operationalized in spatial ability).
- The model contains a possibility to 'loop', because often information-seekers start over within a search process when they cannot find relevant information.
- 'Information scent' is an important and interesting concept that is addressed with this model. The operational definition of information scent in research is the assessed semantic relevance of screen objects to users' goals. Information scent is the feeling a user has about being close to his target or not.
- The model can be used with both searching with a search engine and with browsing categories.
- Because of the fact that CoLiDeS was developed at a more concrete level of abstraction than SNIF-ACT and because CoLiDeS is a semantical model (i.e. the model simulates the semantical similiarity between a description and the initial goal by

calculating LSA-scores), we thought that this model was more appropriate to guide the observations in our research than SNIF-ACT.

In order to clarify what should be represented in cognitive models of information problem-solving, Lazonder and Rouet (2008) differentiated three sets of variables that may affect children's activities during information seeking. As shown in Figure 2.7, these sets include contextual variables, resources variables, and individual variables such as the child's prior skills and knowledge. Lazonder and Rouet (2008) call these three sets of variables 'dimensions of the Information Problem Solving (IPS) activity'. Figure 2.7 presents the IPS dimensions with the IPS activity (i.e. 'access and interaction dimension' of Ingwersen & Järvelin, 2005) in the center of the dimensions.

The steps within the IPS activity are described by Lazonder and Rouet (2008) as follows:

1. Understanding the problem at hand (defining the problem).
2. Using content representation tools by quick scanning or in depth studying.
3. Assessing relevance of information.
4. Transforming information (interpret and present information).

However, the phase of 'using content representation tools' - such as children's informational websites - is not further studied in detail in their research.

Following these dimensions of the IPS activity of Lazonder and Rouet (2008), in the next section, we will discuss prior research on these three sets of variables that may affect children's activities during information searching on informational websites. This

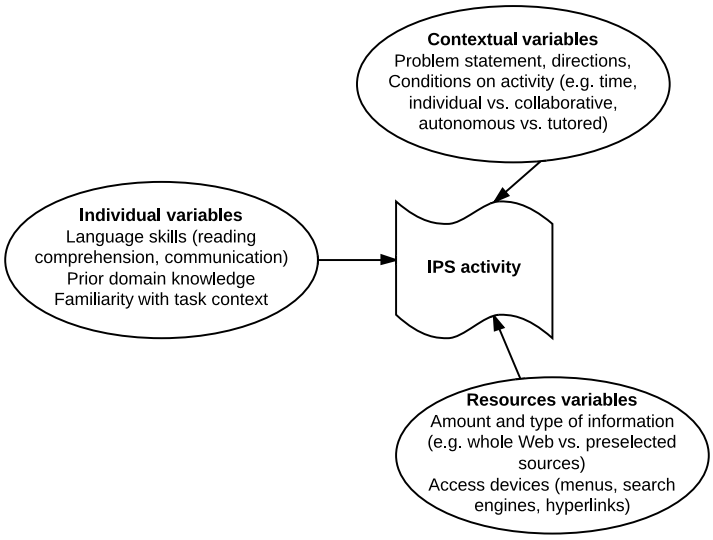


Figure 2.7: Dimensions of Information Problem Solving activity (Lazonder & Rouet, 2008)

will help us in positioning and presenting our own research objectives in relation to these dimensions of IPS activity in the final section of this chapter.

2.2 Which factors influence children's search behaviour?

In this chapter, we will discuss prior research on children's information search behaviour following the dimensions of the IPS activity of Lazonder and Rouet (2008). The overview of different types of research conducted in the field so far supports the choices that we made for our own research objectives. Based on prior research, we argue what domains or which perspectives on children's information search are analyzed, tested or manipulated in our research. Each of the three dimensions of the IPS activity will be discussed in relation to the IPS activity, because in our opinion, it is not useful to discuss contextual or individual variables without discussing it in light of children's actual search activity.

2.2.1 Contextual variables

The first dimension that we want to discuss, refers to the context the user (i.e. child) is in and the problem statement that needs to be solved by the child (see Figure 2.8). Context factors may signal a user to a problem or 'information gap' in his knowledge. For example, a school teacher may give a child an assignment, which is inherent to the educational context of a school. Tanni (2008) conducted research on different types of teaching methods (lectures vs. problem-based curriculum) as a way to influence the context, and found that a problem based curriculum gives students more instruction in critical thinking and judgement of information sources.

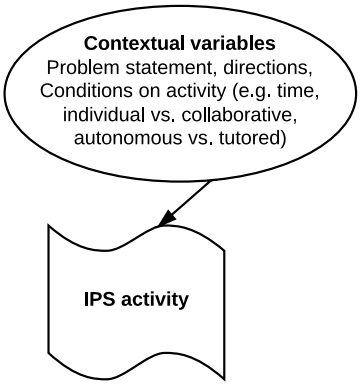


Figure 2.8: Contextual variables of IPS activity (Lazonder & Rouet, 2008)

The context can be of influence on the problem or task, on the user and on the user's interaction with the system. Ingwersen and Järvelin (2005) speak of the social, organizational and cultural context that decides on the domains, goals and work task situations for the actor (i.e. in this dissertation: child). These social, organizational and cultural contexts can also influence strategies, preferences and interests that children have concerning a search task.

General characteristics of context that can be of influence on children's information seeking are social norms and values, purposes and goals, constraints, available information, available resources, amount of distraction and the physical location (e.g. school, library, home, hospital). And in particular for a school setting characteristics that can be of influence are teaching methods (lectures versus problem-based), means and time a child is given, collaboration with other children, or support (with or without instruction). For example, Lazonder and Rouet (2008) reported that collaboration with other students increases awareness of metacognitivity and encourages students to engage in spontaneous reflection. Children trust and obey authority but can be disproportionately influenced by peers.

There are many different approaches towards the problem dimension in prior research. The cognitive view on information needs defines these as follows: 'needs arise from an individual recognizing some dissatisfaction with their existing situation'. Other terms are 'extrinsic uncertainty' (Atkin, 1973) or 'information need as cognitive gap' (Kari, 1998; in Tanni, 2008). Belkin (1982; in Tanni, 2008) sees information seeking as 'mapping texts onto user's anomalous state of knowledge' (ASK). Other approaches to 'information needs' in prior research are 'problems', 'learning tasks' or 'learning assignments' (Tanni, 2008), 'needs (e.g. what someone ought to have for his work) versus wants (e.g. recreational needs)' (Derr, 1983), 'imposed tasks (e.g. by teacher) (Fitzgerald, 2005) versus self-generated tasks (e.g. informal learning or 'just curious' questions)' (Silverstein, 2005), and 'felt needs (individual is aware of his need) versus unfelt needs (individual needs information provider to become aware of his need)' (Kari, 1998).

In the characteristics of a 'problem' or 'task' several variables can be distinguished. For example, Jonassen (2000) describes that a problem can be ill-structured or well-structured, complex or simple, abstract or situated, and general or specific. Cooper (2002) describes problems as well-defined, semi-defined or ill-defined. Bilal (2000, 2001, 2002) conducted research on the differences between different types of tasks: fact-based, research-based and fully self-generated search tasks.

Finally, the goal of an assignment can differ. The goal can be to learn or to select information (Belkin, 1993). The same variables can be named 'researching (understand nature of task)' versus 'reporting (seeking other peoples answers to someone else's question)' (Tanni, 2008). Fitzgerald (2005) mentions the following goals of search tasks: entertainment, fact collection, simple curiosity or collect information to inform consumer

decisions. All these mentioned task variables can play a more or less important role in the information search process.

Defining information need

Prior research has already shown some important findings concerning children's information needs. For example, research reported that children find it difficult to determine what kind of information they need. They have low understanding of their own needs. That is why it is difficult to conduct research on their needs. Also activating their prior knowledge about the subject is hard for children (Walraven et al., 2008).

Conceptualizing information need

Prior research does not really mention the phase of conceptualization of the information need. It is not clear how a child translates an information need to a query or what kind of concepts a child has in mind. Children have difficulties with conceptualizing abstract concepts. That is why they have difficulties understanding abstract headings on web pages. Young children (below age 7) are able to categorize concepts on the basic level (e.g. dog), but logic of classes and taxonomic classification (e.g. animal – dog – Terrier) come at a later age (above age 7) (Siegler, 1991).

Self-generated vs. imposed tasks

In research on informal learning, Silverstein (2005) showed that when late elementary school students are encouraged to bring their own experiences and prior knowledge to the teaching setting, they are more motivated to pursue self-initiated learning. Bilal (2002) reported that children are more successful on fully self-generated tasks compared to imposed tasks. Children also preferred self-generated tasks over imposed tasks.

Search systems should provide children with information that corresponds to their information need. To identify children's information needs mostly sociological methods are used, such as questionnaires or interviews. However, more reliable and precise results can be achieved by using logfiles of search sessions to identify children's query intent. Duarte Torres, Hiemstra and Serdyukov (2010) tried to identify children's queries in a large-scale AOL query log. All queries where the user selected a search result whose domain is listed in the DMOZ's kids&teens directory were regarded as child queries. However, the probability that such pages were also accessed by adults is very high, which makes it difficult to draw strong conclusions from such a study about children's information needs.

Focus of the context dimension in our research

We chose to limit the 'context' in our research to the informational / educational context, because this is the most natural and common context for children in which they search

for information to solve a problem or task. Effects of characteristics within our scope that we considered interesting to examine, were peer-to-peer collaboration, support (with or without assistance or supervision), and effects of different kind of teaching methods. However, the focus of our research was not on potential effects of differences in context factors and we decided to study one 'fixed' informational / educational context in our research, as far as a context can be 'controlled' in a research project. However, as will be discussed in Chapters 4 and 5 of this dissertation, 'support' (the amount of help offered during task performance) proved to be a factor that is difficult to control in research with children.

School setting as physical location

The relevant physical context to examine children's information search behaviour in our research was a school setting, with which we mean that the children used a search interface at school (in the classroom, in a computer lab or in the school library). However, we do think that our research results are also relevant for children's information search behaviour at home when doing their homework for school. Purposes and goals, constraints, available information, available resources, amount of distraction and the physical location can easily be controlled in a research context. We did not manipulate these contextual factors in our research.

Information searching as main activity

The main activity that we examined with children in a digital environment was *information searching*. Activities such as reading, writing stories, watching movies, listening to something, meeting others or making an account/profile were beyond our research scope.

Kind of task

The kind of task that we examined in our research was an *imposed, fact-based assignment*. Of course, this was automatically a learning assignment, but our main focus was not on the learning process. Our main focus was on the process of information searching. Thus, our focus was on children's 'need-to-know' instead of children's 'nice-to-know' and their 'need-to-know' was 'school required'. Although, previous research reports that children are more motivated and successful with fully self-generated tasks, research with these type of tasks would be hard to control, because the content of self-generated tasks would widely differ between children.

Characteristics of task

The characteristics of the task can vary from structured versus ill-structured, to complex versus simple, to general versus specific. The focus of our research was not on effects

of different search task characteristics on children's search behaviour. However, as we will discuss in Chapter 5, we did vary the complexity of the fact-based tasks in our final experiment so that our research results can be generalized over both simple and complex search tasks. In our view, this is more representative for children's search behaviour in general.

Goal of interaction

The main goal of children's information interaction with a search interface in our research was to conduct *research* on an assignment and to *collect* relevant information. We did not want them to reproduce only relevant information that they found, we also wanted them to try to apply this information to the assignment. Further *reporting* on the assignment is beyond the scope of our research.

Types of information

Types of information that was included in our research included mainly school-related subject information (e.g. science, history or arts), information about health, or information, for example, about 'how to take care of my pet'.

2.2.2 Individual variables

The user dimension can be divided into all kinds of characteristics of a child that may be of influence on the child's search process (see Figure 2.9). We divide these user variables in demographical characteristics (e.g. age or gender), cognitive characteristics (e.g. reading ability), affective characteristics (e.g. experienced emotions), and physical characteristics (e.g. motor skills). In this section, we will describe prior research on the effects of different user variables on children's search behaviour. And more specifically, we will describe the user variables that we examined.

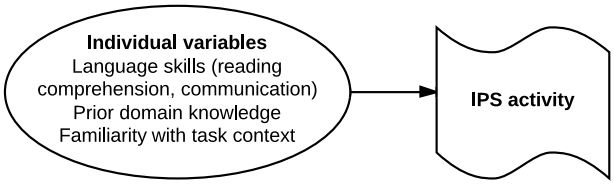


Figure 2.9: Individual variables of IPS activity (Lazonder & Rouet, 2008)

Demographical user characteristics

The age of children is the only demographical user characteristic tested for differences between children in our research. Children aged between 8 and 12 were examined, because previous research shows interesting differences between the skills of children between this age range. For example, the difference between younger children (8-9

years old) and older children (10-12 years old) in conceptualizing is that younger children have concrete, sensory, perceptual, holistic, thematic and global concepts in mind while older children are able to have abstract, symbolic, conceptual, analytic, taxonomic and specific concepts in mind (Siegler, 1991). However, this abstract conceptualization is in development during children's concrete operational stage of development (that starts at age 7) and is usually only fully developed at age 12. This difference in conceptualizing might have consequences for children's formulation of search queries or for children's selection of categories in a navigation structure.

Other demographical user characteristics were explored to check whether they correlated with results found in the experiments:

- *gender*: the results found interesting differences between boys and girls of the same age in interaction with a system.
- *grade-level*: whether differences between grade-levels correlate highly with differences between age groups.
- *cultural background:* we examined children from different social-economical and cultural backgrounds, which depends on the different schools at which our research was conducted, but all children were living in the Netherlands.

Cognitive states and skills

According to Piaget's Theory of Development (Siegler, 1991) children go through different phases of development during their childhood. This means that cognitive states and skills differ between different age groups of children. The following skills can be measured to see whether differences in these skills relate to differences in children's search performance.

- *Reading ability:* reading level might correlate positively with digital literacy.
- *Domain knowledge (factual or structural knowledge):* children are 'universal novices', they lack domain knowledge (Fitzgerald, 2005). Learners with prior knowledge are better able to formulate their own questions, identify resources and choose search terms (Tanni, 2008).
- *Media-expertise:* more expertise in media use or internet experience might correlate positively with digital literacy.
- *Conceptualization skills:* children have difficulties with conceptualizing abstract terms, which has influence on their understanding of navigation headings (Bilal & Wang, 2005).
- *Cognitive learning styles:* children can have a surface, deep or strategic approach; aiming for finishing a task with the least of effort (surface approach) or aiming for personal understanding (deep approach) (Tanni, 2008). These differences in learning styles might also have influence on the way they search for information.

Web literacy

Whether the process of information search is successful may depend on the users' web literacy. Different studies have been conducted on web literacy. Sorapure, Inglesby and Yatchisin (1998) reported that highly web literate users can select relevant information from large numbers of information units and are able to evaluate non-textual elements. Ahtikari and Eronen (2004) approached web literacy from a social-constructivism point of view. They report that highly web literate users should be proficient in the following three areas: content knowledge, skills and strategies and metacognitive knowledge, which are user characteristics that we discussed in Section 2.2.2.

Where Ahtikari and Eronen (2004) describe web literacy in terms of characteristics of the user, Kuiper (2007) describes web literacy in terms of different phases of the search process. She distinguishes the following three areas of web literacy: the search for information, the reading and interpreting of information and the evaluation of information. However, within these areas, she also distinguishes user characteristics such as knowledge of the search engine (i.e. content knowledge), strategies for navigation (i.e. strategies and skills) and reflecting on one's own comprehension of a text (i.e. metacognitive knowledge).

Van Deursen (2009) defines digital skills in operational, formal, information and strategic skills. The main purpose of his study was to identify individual skill related problems that users experience when navigating the Internet. For example, he found that younger subjects (i.e. 18-29 years old) experienced fewer operational and formal skill related problems, but there was no difference regarding information and strategic skill related problems. However, he did not study the digital skills of (young) children.

Metacognitive skills

According to Lazonder and Rouet (2008), metacognitive skills are the ability to plan, monitor and evaluate (i.e. relevance, reliability and authority) one's own actions. A term often used for these metacognitive skills is self-regulation (i.e. thinking about thinking). Examples of self-regulation are goal setting, planning, self-motivation, attention control, application of learning strategies, self-monitoring, self-evaluation, and self-reflection. One's metacognitive skills improve with age and therefore are still in development for children in the age group of our research (Brown & DeLoache, 1978). The ability to monitor and evaluate one's own actions might influence children's search behaviour as we describe in Chapter 3.

Affective mental states

In this dissertation, affective states are discussed in terms of the effects of interface design and search performance on children's affective states. However, affective mental states might also influence children's search behaviour in advance. For example, search

behaviour might be influenced by whether a child is intrinsically motivated to find relevant information, or whether this motivation is extrinsic, for example, because the search task is assigned by a school teacher.

Emotions and attitudes might also influence children's search behaviour in advance. For example, children that are insecure might have more difficulty with searching for relevant information than children that are confident and vigorous about their own information search skills. Children that are averse to searching information, will also experience more troubles with finding relevant information, than children who love to search for information on a website.

Finally, locus of control is an interesting factor. Locus of control is about the personality of the information-seeker and the extent to which he ascribes his failure or success in information search to himself or a system (or another 'locus'). Often the attitude towards a system correlates to a person's 'locus of control'. Unsuccessful information-seekers often do not blame the system, but their own incompetence. This was an important factor to examine with children. Would children blame themselves quickly instead of blaming the system that was not suited for their cognitive abilities and skills?

Physical states: motor skills

Children's fine motor control is not fully developed in comparison to that of adults (Bruckman, 2008). Therefore, children can experience problems during information search concerning motor skills, because often they have to work with devices that have been designed for adults. They can have difficulties using a mouse, because they process information more slowly than adults (Hutchinson et al, 2005). They also have trouble holding down the mouse button for a longer period. Typing can also be difficult for them, because they have to 'hunt and peck' on the keyboard for the correct keys (Borgman et al., 1995). Children's performance with mice and other input devices do increase with age (Hourcade, Druin, Sherman, Bederson, Revelle, Campbell, Ochs & Weinstein, 2002). The more intuitive interaction mode of a tablet computer (e.g. an iPad) might be a solution for many of these problems with motor skills.

Focus of the user dimension in our research

The main focus of our research was not on individual characteristics of the children. In other words, we did not control or manipulate certain individual characteristics to test possible effects of particular characteristics on children's search performance. We did, however, study the effects of children's cognitive skills (e.g. reading level, domain knowledge or media expertise), metacognitive skills, affective mental states and motor skills on their search performance when studying the research results after conducting the different experiments.

2.2.3 Resources variables

The third dimension of the IPS activity (Lazonder & Rouet, 2008) is the system dimension (see Figure 2.10). In this section, characteristics of this dimension are described that can be of influence on children's information search behaviour according to prior research. While the research described was about children's search behaviour (i.e. the IPS activity), the focus in the studies described below is on the role of the system characteristics in the search process.

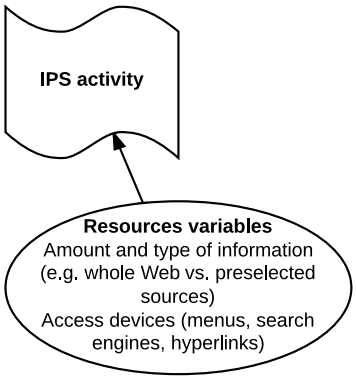


Figure 2.10: Resources variables of IPS activity (Lazonder & Rouet, 2008)

Structure and organization of information

Most research concerning information structures in digital environments for children has examined the difference between children's searching and browsing performance. Many researchers report that browsing categories is more suited for children than searching with a search engine, which we will discuss in Section 2.2.4 in more detail. According to Lazonder and Rouet (2008) for both searching and browsing it applies that children are overwhelmed by large amounts of information and lack of structure in menus or search engines and that children benefit more from carefully constructed resources (e.g. catalogs) than from uncontrolled resources (e.g. Google).

As research claims browsing categories to be more suited for children than searching (Borgman et al., 1995; Schacter et al., 1998), many researchers have addressed different types of navigation structures that children can use for browsing categories. For example, Rose, Rose and Blodgett (2009) compared children's search performance using a map compared to a content list and found that age matters in this comparison. Younger children (between the ages of 7 and 9) recalled more content when using a map compared to a content list, while older children (between the ages of 10 and 12) performed equally with either a map or a content list.

Bilal and Wang (2005) report that the deeper the hierarchies become, the more children become lost. Hutchinson, Druin, Bederson, Reuter, Rose and Weeks (2005) also confirm that broad, shallow hierarchies are better than deep, narrow hierarchies.

In developing the International Children's Digital Library (ICDL), Hutchinson, Bederson and Druin (2006) chose a simultaneous presentation of categories instead of a sequential presentation of categories (see Figure 2.11), because with a simultaneous presentation children can open multiple categories simultaneously. This was found to be faster than a sequential presentation of categories.

Navigation tools

In a study on guidelines for educational websites for children, Meloncon, Haynes, Varelmann and Groh (2010) argue that simple navigation is key to usability in website design for children. Therefore, they recommend limiting navigational topics. "When children are given too many topics, they are easily distracted and often lose their way when trying to use the back-button to return to a familiar place" (Meloncon *et al.* 2010, p. 404).

Further, Meloncon *et al.* (2010) recommend using a visual map for navigation (i.e. image map) instead of simply listing topics in a traditional navigation style. "Image maps enhance visual display of information and decrease the sense of disorientation by information overload", they say. Metaphors that are often used in these image maps should be based on familiar mental models, that are related to age and culture (Large & Beheshti, 2005). According to Nielsen and Gilutz (2002), children like geographical, navigational metaphors of rooms and villages and other simulated environments. Large

Figure 2.11: Simultaneous presentation of categories on the International Children's Digital Library (http://www.childrenslibrary.org/)

(2004) also confirms that children prefer visual cues while browsing for relevant information.

Based on a usability test, Meloncon *et al.* (2010) recommend not including three navigation features often found in websites designed for adults: search options, in-text links and pop-up windows. In particular the first recommendation to not include a search option in a website for children amazes us, because in our opinion, search engines are indispensable for the children of the Google generation.

Content issues

Content for children should be individualized and relevant (Harbeck and Sherman, 1999) and age appropriate and readable (Meloncon et al., 2010). For example, children find 'big words' difficult, that is words of three or more syllables. Also vocabulary needs to be appropriate for the children (Meloncon et al., 2010).

To make content appropriate for children, studies are conducted in which complex language is simplified for children. De Belder, Deschacht and Moens (2010) proposed lexical and syntactic methods of text simplication. In the lexical case they simplified the text by replacing each word one by one with a synonym by using language models from a large, unlabeled training corpus or WordNet, which should be easier to understand. In the syntactic case they split complex sentences into several simple sentences. The researchers did not succeed in reducing the reading difficulty enough for children, at least not without removing information from the text. The syntactic simplification failed. Lexical simplification showed promising results (De Belder, Deschacht & Moens (2010) in Gossen & Nürnberger, 2013).

Other researchers suggest using the complexity of the language in a document to influence an information retrieval system's ranking of the results. Where two topical articles are available, the simpler one should be preferred (Sluis & Van den Broek, 2010). Sluis and Van Dijk (2010) discuss three components that should be considered by an information retrieval system for children: complexity, interestingness and affective value (Gossen & Nürnberger, 2013).

Nielsen and Gilutz (2002) and Naidu (2005) both concluded that children do not scroll down web pages. Nielsen and Gilutz (2002) recommend minimizing the amount of text on screens for children. They report that long pages of text frightened children away, even if they were interested in the content. However, in contrast, Meloncon *et al.* (2010) did not find this to be true in their usability study with children. The children in their study were not averse to scrolling.

When it comes to online help, researchers do not agree. Large and Beheshti (2005) report that children do not read online documentation, which is not the kind of help they need. They want to know for example how to turn an unsuccessful search into a successful one, which is not the help they receive from online documentation (Large &

Beheshti, 2005). Nielsen and Gilutz (2002), however, says children are willing to read instructions in contrast to adults. We do not know the explanation for the different research results between these studies. It might be the case that children became more impatient about reading instructions in the course of the years between the studies.

Interface appearance

Meloncon *et al.* (2010) cite recommendations that children's websites should have an appearance that is simple, clean and concrete, yet also active and enjoyable (Harbeck and Sherman, 1999). Therefore, they recommend web designers to pay attention to graphics, color, games and accessibility. Bruckman (2008) reports that children enjoy a playful approach, which means that graphics that children can relate to should be used.

Large, Beheshti and Rahman (2002) report that children repond positively towards pictures and that children are critical about websites that do not present pictures. However, they also say that children appreciate pictures when they add value to the content. Also animation is an important contribution to a playful approach, because motion is an extraordinarily compelling visual quality that attracts the user's attention (Williams, 2000). Finally, children do not like white, empty space and they prefer bold, vivid and vibrant colors (Large and Beheshti, 2005; Meloncon et al., 2010).

Operational devices

The desktop computers that have been used for years now with the mouse and keyboard as input devices are not very intuitive for children. These computers have not been designed with the interactions of children in mind or with the goal of supporting learning (Inkpen, 2001). Touch screen technology is a more natural input device than the traditional input devices. The most important advantage of touch-screens is that the input device, is also the output device. There is a sense of immersion, because the user is able to touch, feel and manipulate objects on a screen (Holzinger, 2003). Therefore, this might be a more effective input device for children. A disadvantage of touch screens is that the current tools are not yet optimally aligned for conducting complex tasks on a screen.

However, with touch screens, the user is also interacting with a screen. This is just a window through which digital information can be reached, as is the case with all graphical user interfaces (GUI's). A real sense of immersion by touching, feeling and manipulating objects, is when they are not on a screen but when they are real tangible objects (TUI's; tangible user interfaces). This type of direct interaction would be the most 'natural' form for children, because children are born with the ability to tangibly manipulate objects (Xu et al., 2007).

Focus of system/interface characteristics within our research

Of course, it is not feasible to study all these mentioned system and interface characteristics in one research project. Therefore, to decide what characteristics are most relevant to study in our research project, we conducted both a corpus study to analyze the characteristics of children's informational interfaces and an explorative study of children's interaction with existing informational interfaces. These two studies led to the design of an experiment in which several interface characteristics were manipulated. Of the characteristics mentioned in this section, we will study and discuss the following concerning children's search behaviour in the different chapters of this dissertation:

- Structured and carefully constructed resources (i.e. informational websites for children) versus uncontrolled resources (i.e. Google) (Chapters 4)
- Browsing categories versus keyword searching using a search engine (Chapters 4 & 5)
- Text-based retrieval/menus versus image-based (graphical) retrieval/menus (Chapters 4 & 5)
- Textual versus visual categories (Chapter 4 & 5)
- Menus versus graphical metaphors (Chapters 4 & 5)
- Playful versus classical design approach (Chapters 4 & 5)

2.2.4 Interaction dimension

The interaction dimension is the core of the IPS activity (Lazonder & Rouet, 2008) as presented in Figure 2.5. The information search process revolves about the interaction between the context, the user and the system. Besides topical well-defined requests on content (i.e. keyword searching), Ingwersen and Järvelin (2005) also mention other kinds of strategies of information seeking, such as browsing and navigation. In this section, we will discuss previous research concerning interaction between the child and the system, such as children's search strategies and navigation behaviour on informational websites. Where the focus of the studies discussed in the previous section was on the system characteristics, the focus in this section is on the actual behaviour (e.g. searching, browsing and evaluating information), performance (e.g. effectiveness and efficiency of conducted search tasks) and system evaluation (e.g. satisfaction with, attitude towards and evaluation of the system/interface) of the children during the information search process.

Query formulation

Prior research reported the following problems that children's experience during query formulation. Children have difficulties with formulating queries, because they have less knowledge to base 'recall' on than adults (Borgman et al., 1995), they rarely access their previous knowledge during query formulation (Schacter et al., 1998) and they have less

vocabulary knowledge than adults (Borgman et al., 1995). The most important problem for children's querying is that they have difficulties with correct spelling, spacing and punctuation. They find it difficult to transpose natural language into a single keyword. They have difficulties with reformulating unsuccessful queries, because they do not know many synonyms and they find it difficult to specify search queries.

McCrory Wallace, Kupperman and Krajcik (2000) found that students had trouble focusing on and refining their research questions. Instead of focusing on content and refining questions based on finding information, they reacted on irrelevant information by proposing to change their initial question. Instead of focusing they moved the lens from place to place.

Exploration by navigating

This stage of the interaction describes the actual exploration of information; planning search strategies, locating sources and searching relevant information. Several information searching patterns can be identified: fast surfing, broad scanning and deep diving. The information search pattern of 'fast surfing' is often found by children. They often search documents for the 'right' answer instead of constructing their personal understanding (Tanni, 2008).

Marchionini (2006) identified the following search strategies: lookup, learn and investigate. Lookup search is also called 'fact retrieval' or 'question answering'. Lookup is the most basic search task and has been the focus of development of information retrieval systems. However, as the web became the resource of first choice of information seekers, people expect more from information systems. Systems have to support their exploratory search to 'learn' and 'investigate' beyond just retrieving facts.

Children search for the 'exact match' and for the 'right answer', not for understanding the information. A lot of children overemphasized the search aspect of the process, treating searching itself as the centerpiece of the work they did, and not 'finding the right information' or understanding and learning from this information (McCrory-Wallace et al., 2000).

Prior research reported the following characteristics of children's information search behaviour. Children are interactive information-seekers who do not tend to preplan their searches (Schacter et al., 1998). They are reactive searchers; they do not use sophisticated analytical searching strategies. They search by excessive browsing (Chen, 2003). Children do not save previous searches and they often repeat unsuccessful searches (i.e. loopy-style) (Hirsh, 1999).

According to Bar-Ilan and Belous (2007), browsing directory structures is better suited for children's information retrieval, but children prefer keyword searching. Cooper (2002) stated that browsing can also be more difficult for children than for adults. Children have difficulties in understanding and selecting abstract terms; search tasks

are more successful when concrete terms are used and children find it easier to retrieve concrete terms than abstract terms. Children can also have trouble understanding categories and finding the right category, because they have less domain knowledge and less vocabulary knowledge than adults (Cooper, 2002).

Children do not record URL's or websites. Instead they try to recreate successful searches. They prefer to use the back button to return to good information than to bookmark relevant results. They do not store information or elaborate on content and they use relevant information immediately as they see it.

Children rarely surf more than a few minutes on a website, they click and move from place to place and they seldom read beyond the first screen (Chen, 2003). They rarely scroll pages and mainly interact with information that is above the fold (Nielsen, 2002). Instead of surfing innumerable pages to explore available information, McCrory Wallace *et al.* (2000) found that students stayed very close to home. They stayed within five hyperlinks from their initial list of search results.

Children are more chaotic in their search performance than adults. They make more web moves than adults, backtrack more often, loop searches and hyperlinks more often and deviate more often from a designated target (Chen, 2003). Bilal and Kirby (2002) concluded in their research that adults adopt a 'linear or systematic' browsing style whereas most children have a 'loopy' style. The same difference between the browsing style of adults and children was also reported by Gossen, Low and Nürnberger (2011).

Information collection, use and evaluation

This stage of the information search process describes children's behaviour when processing found information and gathering, using and interpreting this information to construct new knowledge and formulate an answer to the initial question.

The strategy often found by children to process information is to linearly scan documents to search for the exact match to their query. While scanning web pages, they mainly scan pictures and first paragraphs of web pages. Young children rarely read web pages in-depth. They do not store relevant information, but modify text in their own words. They do not read text to understand in depth (Walraven et al., 2008).

For example, Brown and De Loache (1978) describe the task of visual scanning. Young children are not yet able to coordinate and control their scanning activities, but children's scanning skills gradually become more systematic. The older the child, the more likely he will fixate on features that are relevant to his goal. Younger children have difficulties ignoring irrelevant features. The younger the child, the more attention he is likely to give to stimuli that are irrelevant to the original task. The implication for their interaction with digital environments may be that children will have more difficulties to parse a web page into subregions and to focus their attention on relevant items, if the screen is filled with irrelevant visual features.

Elementary school children find it difficult to evaluate information. Most children do not even evaluate information on the web or are not interested in evaluating information at all. They do not have critical thinking skills to evaluate the authority of websites. However, the ability to evaluate information increases with age and education (Dunn, 2002). Prior studies report the following findings on children's search and relevance criteria to evaluate information:

- Children use different search criteria than adults. They prefer non-traditional search criteria such as colours, shapes, genres or feelings and they cannot make a clear distinction between fiction and non-fiction (Druin, 2003).
- The most important relevance criteria for children are topicality/aboutness and concrete/exact answer (Bilal, 2000). Also novelty and interest are mentioned as important relevance criteria for children by Walraven *et al.* (2008).
- Children look at covers, titles and pictures as criteria to select information. Their judgment is based on appearance, length of text and use of language. Some children base judgment on rank in hitlist or on number of results their search produced. (Walraven et al., 2008)
- Young children tend to think everything is true on the internet and authority of sources is not considered by young children (Walraven et al., 2008).
- Children rely heavily on summaries and note fields to decide whether information is relevant (Hirsh, 1999).
- Relevance criteria that are not considered by children are: authority, recency, truthfulness, accuracy and validity (Walraven et al., 2008).

Measuring quality of use: performance and evaluation

One of the pioneers in evaluation research of 'quality of use' (Bevan, 1995) introduced two perspectives on quality of use: an interaction-centered perspective (effectiveness and efficiency) and a user-centered perspective (satisfaction). The difference between these two perspectives is also that the first is an objective perspective on quality of use and the second is a subjective perspective on quality of use. This definition of 'quality of use' was also included in the ISO 9241-11 standard: 'Guidance on Usability' (1998). These two perspectives on quality of use correspond to the two perspectives that we used in our study: 'performance' and 'evaluation' as presented in the Introduction of this dissertation.

The research on children's interaction with digital products over the last decade is also dominated by pragmatic issues, such as the utility and usability of technical systems (Thüring & Mahlke, 2007). Mostly, performance-based methods were chosen to assess two usability components: effectiveness and efficiency (Bilal, 2002). A third component of usability that is often studied from a user-centered perspective is user satisfaction. However, subjective judgments of the use of the system to measure user satisfaction

are also mostly based on the efficiency and effectiveness of system usage (Lindgaard & Dudek, 2003). Hassenzahl, Platz, Burmester and Lehner (2000) say: "We are aware that user satisfaction is a part of the usability concept provided by ISO 9241-11. However, it seems as if satisfaction is conceived as a consequence of user experienced effectiveness and efficiency rather than a design goal in itself. This implies that assuring efficiency and effectiveness alone guarantees user satisfaction."

Pragmatic variables that are mostly measured with these performance-based research methods in research with children are search success (Rose, 2009), search times (Inkpen, 2001), Web moves (e.g. searching, browsing, looping, backtracking, scrolling, mouse movements, exploratory moves, or selection moves (Bilal, 2000), search efficiency (Bilal, 2000), error rate (Inkpen, 2001), ability to construct a search query, information recall (Rose, 2009), "weighted" traversal effectiveness (Bilal, 2000), search method consistency behaviour (Hirsh, 1997), difficulty with using the interface (Hutchinson, 2006), comprehension of search task (Hutchinson, 2006), number of keywords, visited pages, depth of exploration, average exploration depth for task completion, revisited pages, and frequency of refining or improving answer quality (Kao, Lei & Sun, 2007).

Usability is not all that matters in human-technology interaction. It is also important how children feel during task performance and how they emotionally respond to a system (Thüring & Mahlke, 2007). In 2002, Dillon already proposed that user satisfaction is likely to be influenced by other factors such as personal experience with technology and the aesthetics of system design. The field of *User Experience* (Hassenzahl & Tractinsky, 2006) argues for a broader perspective of user experience than only usability. Besides instrumental qualities (i.e. pragmatic quality / usability), also non-instrumental qualities are important, such as visual aesthetics or haptic qualities. Furthermore, it is important to know how people feel while using a system. These new ideas regarding the subjective user-centered perspective on quality of use (i.e. our 'evaluation' perspective) were already seen in research on fun of use (Carrol & Thomas, 1988), emotional usability (Logan, 1994; Kim & Moon, 1998), ludic products (Gaver & Martin, 2000), pleasurable products (Jordan, 2000), hedonic quality (Hassenzahl et al., 2000), visual aesthetics (Lavie & Tracstinsky, 2004), affective quality (Zhang & Li, 2005) and product emotions (Desmet & Hekkert, 2007).

Thüring and Mahlke (2007) developed a framework (see Figure 2.12) to support an experimental approach in research on human-technology interaction, termed the Components of User Experience model. This model elaborates on the subjective user-centered perspective on quality of use (i.e. evaluation perspective) of an interactive product. Three types of User Experience components are identified: perception of instrumental qualities (such as usefulness and usability), emotional user reactions (such as subjective feelings) and perception of non-instrumental qualities (such as aesthetic aspects). All three components mentioned have an influence on the overall evaluation

of the interactive product (as can be seen in Figure 2.12). This framework supports researchers in designing experiments on User Experience for influencing factors that can be varied, for differentiation of components that can be measured and for formulating hypotheses about interrelations of the defined variables.

Examples from research with adults that measure user experience components are measuring perceptions of hedonic qualities and evaluational constructs such as beauty and goodness by means of bipolar verbal anchors (Hassenzahl, 2004; Van Schaik and Ling (2008). Hedonic quality refers to the perceived pleasure-producing qualities of the website to achieve be-goals, such as 'being special', or 'being related to a particular group'. Two forms of hedonic quality are stimulation (e.g. the product is original, creative, innovative or exiting) and identification (e.g. the product is integrating, professional, inclusive or presentable) (Van Schaik and Ling, 2008). Tuch, Roth, Hornbaek, Opwis and Bargas-Avila (2012) measured perceived usability and perceived aesthetics by means of scales and items. Furthermore, they measured emotional valence using the self-assessment manikin (SAM) scale from Lang (1980).

However, not much research on children's search behaviour measures children's user experience variables except for some measurements of likeability (Hutchinson et al.,

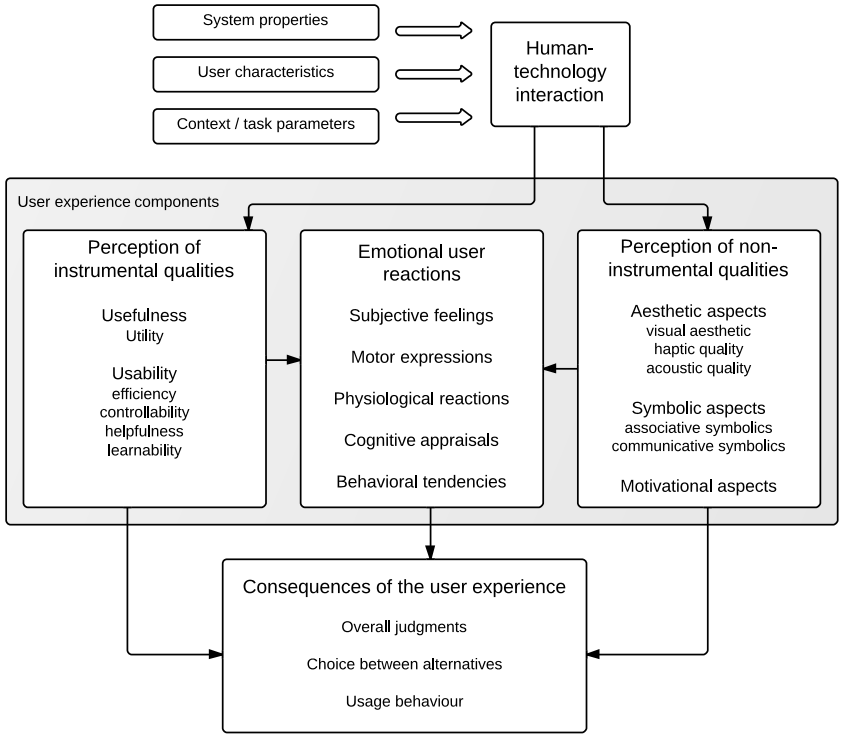


Figure 2.12: Components of User Experience (Thüring and Mahlke, 2007)

2006) and preference (Inkpen, 2001), which is also most likely based on efficiency and effectiveness of system usage. There are some studies on how to measure subjective feelings with children. In a study to gather opinions in child computer interaction, Read and MacFarlane (2006), for example, showed that the major factor in children's decision about whether they want to use an interactive product again is how much fun it was instead of ease of use. They develop, tested and validated the Fun Toolkit to record children's opinions about technology and to gain a measure of children's engagement, existing of the Smileyometer, the Funometer, the Again-again table and the Fun Sorter (Read, MacFarlane & Casey, 2002; Read and MacFarlane, 2006; Sim, MarcFarlane and Read, 2006; Read, 2008). Other developed and validated evaluation tools for evaluating children's technology are the Thumbs-Up Scale (TUS) and Frequency of Use Scale (FUS) (Kano, Horton and Read, 2010) and the Fun Semantic Differential Scales (Yusoff, Ruthven and Landoni, 2011).

Measuring user experience components (such as perceptions of instrumental and non-instrumental qualities and emotional reactions) of a system is difficult, particularly with children, because of risks of satisfying, children's tendency to say yes irrespective of the question and children's tendency to indicate the highest score on the scale when scales are used to elicit opinions about software (Markopoulos, et al., 2008). Horton, Read and Sim (2011) report a study in which the reliability of children's responses on a pictorial questionnaire is tested by asking the same questions about children's technology twice one week after the other. None of the children produced the same results for a question after one week, which shows the difficulty of using survey questions with children and the issues with the validity and reliability of survey answers given by children.

Focus of interaction characteristics within our research

In our research the focus was both on effectiveness and efficiency of children's search performance and on children's evaluation of search interfaces (i.e. perception of instrumental and non-instrumental qualities, emotional reactions and overall evaluations). On the one side, we were interested in the effects of different interface design characteristics on children's objective pragmatic performance, such as the effectiveness of their search tasks and the amount of time and number of clicks they need to find relevant information. On this side, we also studied the successes or problems children experience with keyword searching or browsing through categories. On the other side, we were interested in the effects of these different interface design characteristics on children's subjective perceptions of instrumental and non-instrumental qualities, their affective responses (such as emotional valence and arousal) towards the system during the information search process and their overall evaluation of the system (such as beauty, goodness and fun). The relation between both pragmatic performance measures and measures of user experience components of children's interaction with the interfaces

also perceived our attention. For example, did children like to use a search interface because it was easy to use, or because they thought the interface was beautiful or fun to use?

2.3 Our research objectives

In this chapter, we have given an overview of research traditions that are relevant for our research project and the models that are commonly used in these research traditions to study information search behaviour. Then, we gave an overview of research factors that may have influence on children's information search behaviour. These research results were reported by following the dimensions of the IPS activity of Lazonder and Rouet (2008). For each of these dimensions, we described which variables would be studied in our research project to be able to answer our main research questions. Finally, we presented two perspectives to evaluate quality of use of children's search interfaces: by using a performance perspective (i.e. effectiveness and efficiency of children's search performance) and by using an evaluation perspective (i.e. children's perception of instrumental and non-instrumental qualities, emotional reactions and overall appraisal of a search interface). Figure 2.13 gives a summary of the variables that we studied in our research project, presented in the same manner as the dimensions of the IPS activity of Lazonder and Rouet (2008).

In Chapters 4 and 5, children's performance scores and evaluation scores in interaction with the mentioned contextual, individual and resources variables will be reported. First, by exploring children's behaviour on existing interfaces (Chapters 4) and after that, empirical tests carried out (Chapter 5) by creating three conditions of the same interface based on the results of the corpus study (Chapter 3) and the explorative study (Chapter 4).

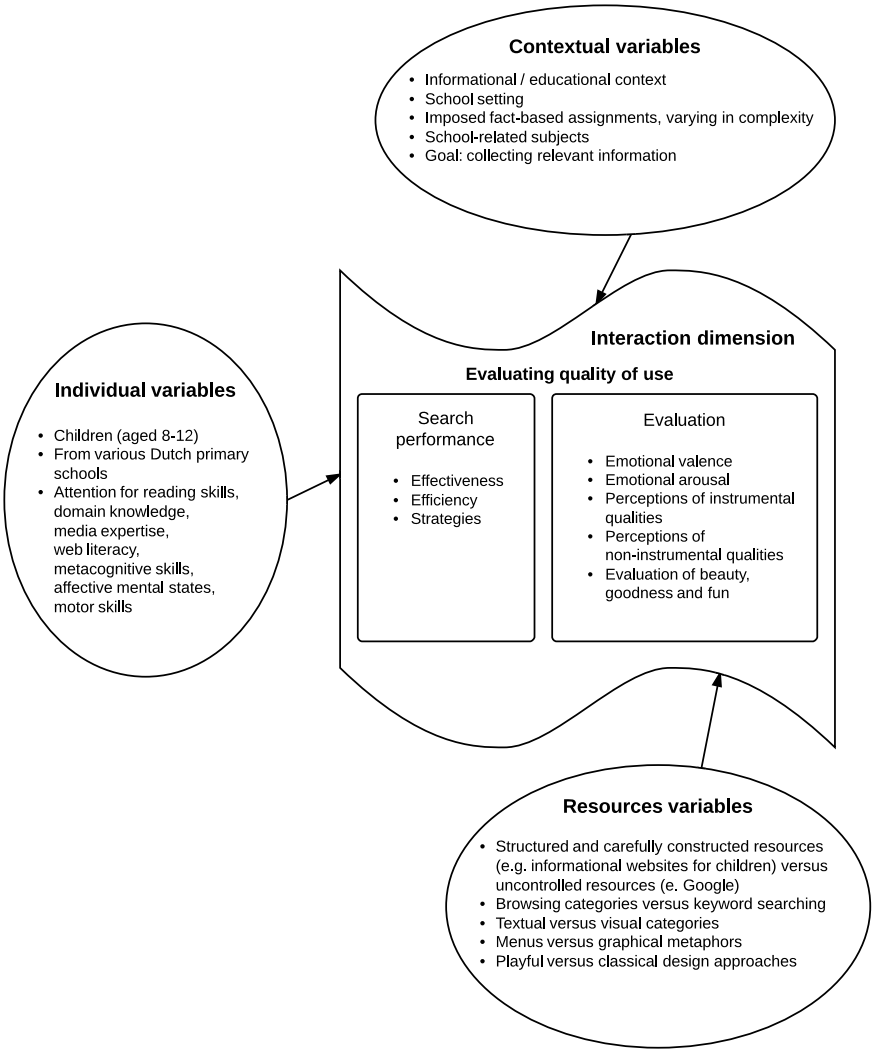
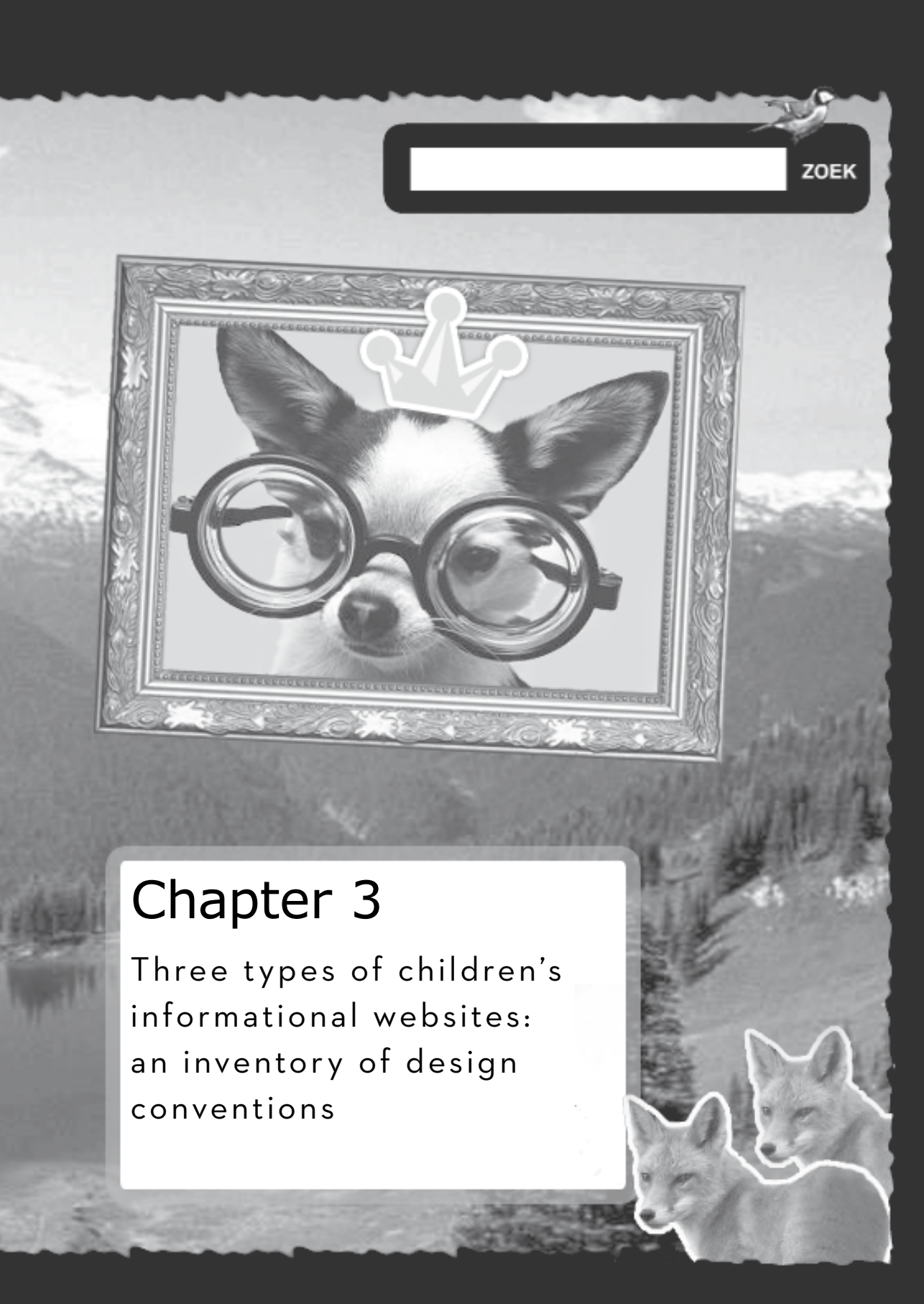


Figure 2.13: Focus of the dimensions of the IPS-activity in our research



Chapter 3

Three types of children's informational websites: an inventory of design conventions

3.1 Introduction

Children are one of the fastest growing groups of Internet users in the United States and Europe (Nielsen Wire, 2009). With this development, the number of websites especially designed for children is also increasing rapidly. Despite the growing body of research on children's search behaviour (Borgman, Hirsh, Walter & Gallagher, 1995; Schacter, Chung & Dorr, 1998; Bilal, 2001; Large & Beheshti, 2002; Druin, 2003; Hutchinson, Druin, Bederson, Reuter, Rose & Weeks, 2005), it is not clear whether design conventions that are used for adult users, are also suited for children's interactions with web interfaces. Before conducting research on the question of what design conventions are suited for children, we first felt the need to identify the current conventions used in informational websites for children. Therefore, in this article the state of the art in children's informational websites is presented.

At the moment, there is no clear set of heuristics that can guide the design and evaluation of children's websites. Web design for children is mostly based on 'myths' about what children prefer and what is easy for them to use. Most of these myths are not based on research, but are a projection of adults' visions about children's preferences. Often, adult content is taken and made 'childlike' by simplifying the content and by adding more visual design and multimedia (Nielsen & Gilutz, 2002).

Nielsen and Gilutz (2002) do propose a set of guidelines for children's web design, but many of the guidelines appear to be quite similar to standard web design practices for adult websites (Markopoulos, Read, MacFarlane & Hoysniemi, 2008). Meloncon, Haynes, Varelmann and Groh (2010) recommend a set of guidelines to direct the design process for children's websites by conducting a usability test of an educational website for children. Because of the few published resources with an exclusive focus on children, they supplemented the available resources about web design for children with basic adult guidelines from Spyridakis (2000) and Krug (2006). The adult guidelines that were used were tested and validated in their study with children.

Users rapidly become accustomed to prevailing design standards and conventions. Therefore, when users arrive at a website, they assume it will work the same as other sites (Nielsen & Loranger, 2006). For this reason, guidelines for web design should have a close relation to the actual practice of web design. However, there is no widely accepted notion of what the genre of children's informational websites for children between 7 and 12 years old looks like. In this article we present the results of a corpus analysis of 100 children's informational websites. Our aim is to make an inventory of the conventions used for children's informational websites. The main question that will be answered in this chapter is *RQ 1* of this dissertation:

RQ 1. *Which design conventions on interfaces for children are specific for children and which of them are general design conventions?*

This question will be answered through a corpus analysis that identifies the current conventions used in informational websites for children.

Besides studying the overall design characteristics from the corpus, a closer look was taken at subsets of the corpus. We discovered three different types of children's informational websites. The design characteristics of these different types will be described in more detail. We will discuss whether these design characteristics are relevant to children's cognitive development and their needs and preferences. We expect these insights to be of value for future research on design characteristics that do or do not support children's interaction with informational websites. Before presenting the results of our corpus analysis, we will first discuss the literature on why children are different thinkers than adults and the implications of children's cognitive development for their interaction with different web design characteristics. The identified web design conventions for children in this study can be addressed in light of these implications.

3.2 Children's cognitive development in relation to design

Early developmental psychologists, such as Piaget (1964), Vygotsky (1962) and Bruner (1966), stated that children are a fundamentally different kind of thinkers than adults. Carey (1988) reports two arguments why children are fundamentally different thinkers according to these early developmental psychologists: 1) Children differ from adults in the kind of concepts they can represent mentally (i.e. conceptualization) and 2) children cannot think about their mental representation (i.e. meta-cognition).

1. Young children are already capable of categorizing concepts on the basic level of abstraction (e.g. dog), but have more difficulties with understanding concepts on superordinate levels with a higher level of abstraction (e.g. animal). Logic of classes and taxonomic classification (e.g. animal – dog – Terrier) come at a later stage above the age of seven with children (Bilal & Wang; Siegler, 1991). For children's interaction with digital environments, this may imply that children have more difficulties browsing through abstract main categories than through main categories on a basic level of abstraction (e.g. children might find it difficult to search for 'chameleons' or 'lizards' in the category named 'reptiles').
2. Later developmental psychologists, such as Siegler (1991), Carey (1988) and Brown and De Loache (1978), argue that by far the most important source of variance for cognitive tasks is in domain-specific knowledge. When faced with a new type of problem, everyone is a novice, both children and adults. Novices often perform inefficiently, because they lack skills, but also because they lack self-conscious participation and self-regulation of their actions (Brown & De Loache, 1978). Children

experience this far more often than adults, because they are 'universal novices'; most tasks they have to conduct are new to them (Brown & De Loache, 1978). While conducting memory or problem-solving tasks, children often fail to check and monitor their activities and fail to make their own task analysis. This monitoring is termed meta-cognition: the voluntary control an individual has over his own cognitive processes. To build up knowledge and to adopt self-regulation strategies, children need experience.

Next, the implications of children's cognitive development will be discussed for their interaction with different types of web design for children, such as scanning, searching and browsing.

3.2.1 Scanning

A relevant task for digital information seeking that is decribed by Brown and De Loache (1978) is the task of visual scanning. Young children are not yet able to coordinate and control their scanning activities, but children's scanning skills gradually become more systematic. The older the child, the more likely he will fixate on features that are relevant to his goal. Younger children have difficulties ignoring irrelevant features. The younger the child, the more attention he is likely to give to stimuli that are irrelevant to the original task. The implication for their interaction with digital environments may be that children will have more difficulties to fixate on relevant items, when the screen is filled with irrelevant visual features. The process of scanning is a natural response that does not develop itself. What develops is the ability to control and coordinate this scanning process to use it as a strategic action (Brown & De Loache, 1978).

3.2.2 Searching

Young children get distracted very easily and can forget their initial search goal while searching. When children grow up, their search strategies become more sophisticated. Their attempts become more systematic and efficient and their strategic use of cues for retrieval improves with age. The implication for searching in digital environments is that younger children especially need clear cues to help them search and that distractions should be limited.

External retrieval (i.e. scanning externally provided categories) is easier than internal retrieval, because internal retrieval (i.e. scanning your own memory), requires a higher degree of cognitive control to initiate and maintain the cognitive orientation to the relevant information in memory. When some kind of structure is provided, which is the case with external retrieval, a child will perform much better than when he must make up that structure for himself (Brown & De Loache, 1978). The implication for searching in a digital environment is that, from a cognitive perspective, children may have more difficulties with keyword searching than adults. For 'retrieving' a relevant query from

memory, the child needs a higher degree of cognitive control and orientation than for 'browsing' through provided categories.

Other reasons why children experience difficulties with keyword search are that children have less knowledge to base 'recall' on than adults (Borgman et al., 1995, Hutchinson et al, 2005) and that they rarely access their previous knowledge of the topic during formulating search queries (Schacter et al., 1998). Probably because they do not know that they have to access this knowledge, because they lack metacognitive skills through their inexperience with these types of tasks. Their vocabulary is not as well developed as an adult vocabulary either and they have difficulties with correct spelling, which is needed for most search engines to find relevant search results. Also, moving from natural language to a single keyword is more difficult for children than for adults (Rowlands, 2008). Still, while keyword searching is more difficult for children than browsing from a cognitive perspective, children seem to prefer keyword search over browsing (Bilal & Watson, 1998; Large, Beheshti & Moukdad, 1999; Druin, Hutchinson, Foss, Hatley, Golub, Leigh Guha & Fails, 2009).

3.2.3 Browsing

Browsing categories should be less difficult for children than keyword search, because some kind of structure (e.g. a row with main categories) is provided. However, browsing can also be more difficult for children than for adults. As already mentioned at the beginning of this section, children have difficulties in understanding and selecting abstract terms; browsing tasks are more successful when concrete categories are used (Borgman et al., 1995) and children find it easier to retrieve concrete terms than abstract terms (Bar-Ilan & Belous, 2007). Again, children can also have trouble understanding categories and finding the right category, because they have less domain knowledge and less vocabulary knowledge than adults (Borgman et al., 1995, Bar-Ilan & Belous, 2007). Also, the number of categories is important. Cognitive load increases with the number of categories that need to be processed. The more categories to choose from, the more difficult it is for a child to choose the right category. Further, we expect children to use a trial-and-error navigation strategy more often than adults, because children's meta-cognition is not fully developed (Brown & DeLoache, 1978). This is because children have less control over their own cognitive processes and less ability to monitor and reflect on their choices.

From the results of the corpus study, we will discuss which aspects of digital interfaces might be supportive for children's cognitive skills and which aspects might cause problems for the search processes discussed in this section such as browsing categories, scanning web pages and searching information using keywords.

3.3 Method

For the corpus analysis, 100 websites were selected to form a representative corpus for children's websites in general. For the website selection and analysis, children's scenarios were developed for informative websites (de Jong & Lentz, 2006). These scenarios were based on four contexts that are common in children's lives. When relevant information concerning these scenarios could be found on or through a website (e.g. through a directory or search engine), we included this website in the corpus. We choose to analyze the websites on the basis of these scenarios and not by looking at the website when filling in the coding scheme (i.e. heuristic analysis), that will be discussed later in this section. Searching for relevant information is more representative for real website use than just 'looking around'.

3.3.1 The corpus

Content

The intended communicative effect of the websites had to be informational (and not advertisement or entertainment). Websites were considered as informational and were selected for our corpus when one of the four scenarios presented in the next section could be performed on the website and relevant information could be found as an answer to the search question. Also search engines or directories were included when relevant information for the scenarios could be found through these search engines or directories.

Audience

Websites were aimed at children in the age category from 7 through 12 years old. Most websites mentioned the age category that their website was aimed at in their 'About us' section. We also included websites that were aimed at a broader age category than 7 through 12 years old (e.g. also for younger children or also for teenagers), because in these cases, the website should also be usable for our age category. Some websites had different sections for different audiences. In that case, only the section for our target age was visited. When a website did not mention the aimed age category, we checked whether words such as 'children', 'child', 'kids', or 'kid' were used on the home page or in the title of the website. When this was the case and the website predominantly contained text instead of images, then we argued that this website was aimed at children in the age from 7 to 12 years old. This is the age category in which children can read and comprehend informative texts.

Information structure

Websites had a minimum information hierarchy of three levels. Websites with one navigation level were not included in the corpus (e.g. the website fjordstone.com/kidzong). The reason for this is that these different levels provide the opportunity to navigate through a website from the home page with the main menu, to a sub page with a sub menu, to a text page with the relevant information. Search engines and directories are an exception in respect to this criterion, because they were also included in our corpus, even though these websites do not have a minimum information hierarchy of three levels. These search engines and directories do not provide actual content on the website; they refer to other websites that do provide the relevant information.

Language

To develop an international corpus, websites were selected in three different languages that did not need to be translated for the analysis. The majority of the websites in the corpus were in English (80 websites in total). Twelve websites were in Dutch and eight in German. The preponderance of English websites in the corpus represents the preponderance of American and British websites in the western world in relation to German and Dutch websites.

Accessibility

Websites that require registration to read content were not included in the corpus. Also, websites that require a download of specific software to view the contents were not included in the corpus.

Source

We chose to select websites from one internet directory, DMOZ.org. This is an open source directory in which volunteers work on building the largest human-edited directory of the web. The directory for children (The Kids and Teens Open Directory Project) has the goal to produce the most comprehensive directory of the web for people under the age of 18, by relying on a vast army of volunteer editors. Within this directory, all websites are selected based on one set of criteria. In Appendix B, an overview can be found of the names of all 100 informational websites from the corpus.

3.3.2 Scenarios

Children operate in several different physical, social and cultural contexts. They spend most of their time at *school* or at *home*. Besides that, they spend time on *leisure*, for example, by playing sports, making music or pursuing other hobbies. Some children may spend time in less common contexts. For instance, when they are ill, they might need to spend some time in a *hospital*. The four contexts mentioned here were selected

because children might use digital interfaces in these contexts to acquire knowledge. By analyzing websites from these four different contexts, we hope to have reached a corpus that is representative for a large group of informational websites for children. We selected 25 websites per context to reach a total of 100 websites in the corpus. To make the website selection per context more specific, a domain was chosen per context to search for information about, namely, science, pets, arts and health (see Table 3.1). For each of these domains, a short scenario was followed on the websites to search for specific information.

Table 3.1: Contexts, domains and scenarios for website selection

Context	Domain	Scenario
School	Science	Jim wants to give a presentation for his classmates about dinosaurs and searches for information on the internet.
Home	Pets	Bob's parents surprised him with a puppy for his birthday. He wants to know how to take care of it and searches for information about puppies on the internet.
Leisure	Arts	Emily loves to make music or listen to music and searches for information about an artist or instrument on the internet
Hospital	Health	Rose has diabetes and wants to know more about her disease. She searches for information about her disease on the internet.

3.3.3 Coding scheme

The corpus analysis was conducted in the spring of 2010 by four trained research assistants who received course credits for their work. Before conducting the final analysis, two pretests were conducted with five and ten websites respectively, in which adjustments were made to the initial coding scheme.

The assistants filled in the coding scheme in a window next to the browser window in which they analyzed the different websites. It took the assistants approximately half an hour per website to fill in the coding scheme. Each of the four assistants analyzed 50 websites from two domains, so that each website in the corpus was analyzed by two different assistants.

Each of the websites was analyzed by following the scenario that was relevant for the domain that the website belonged to. By following a short scenario (see Table 3.1), the assistants were guided to the same web pages per website, which made their evaluations of websites more comparable.

Inspired by the design categories used in literature on existing guidelines for web design for both children and adults (Farkas & Farkas, 2000; Spyridakis, 2000; Williams, 2000; Nielsen & Tahir, 2002; Koyani et al., 2006; Meloncon et al., 2010), we developed a coding scheme for analysis of the websites for children in which we included the main and subcategories mentioned in Appendix A.

For each website, we listed the location or presence of items on the web page (such as the logo, the main menu, the search window, etc.), the scanability and readability of the

web pages and the type of media used (such as images, video and audio). These design characteristics of the websites may have implications for children's scanning behaviour. We also listed the characteristics of the navigation tools and the search engine, which may also have implications for children's searching and browsing behaviour. Within the category 'visual design', three levels of web pages per website were analyzed, the home page, a sub category page and a content page.

3.3.4 Inter-coder reliability

To check whether the results of the four assistants were reliable, we computed the inter-coder reliability using Cohen's kappa, which measures the amount of agreement between assistants. When kappa is above .60, there is a fair agreement between assistants. We measured Cohen's kappa between the scores per domain for the home pages.

As can be seen in Table 3.2, the scores were least reliable in the Arts domain. Therefore, we took a closer look at the scores of the assistants and we found substantial disagreement between the scores of one of the assistants and the scores of the other assistants. Therefore, we decided to resolve all disagreements between assistants per domain by consulting an independent, fifth researcher (i.e. the author of this dissertation).

Table 3.2: Kappa scores for research assistants' reliability

Domain	First researcher	Second researcher	Cohen's kappa	Cronbach's alpha
Science	A	B	.86	.97
Pets	A	C	.67	.89
Arts	D	C	**.51**	.64
Health	D	B	.67	.94

3.3.5 Definition of convention levels

According to Nielsen's (2004) definition of standardization levels, there are not many design 'standards' in our corpus of children's websites, as will be discussed in the Results section. According to Nielsen (2004) a 'standard' is when 80% or more of the websites use the same design approach. When 50-79% of the websites use the same design approach, Nielsen speaks of a 'convention'. When fewer websites use a particular design approach, he speaks of confusion.

For the results of our corpus analysis, we decided not to speak of 'standards', because the analyzed design characteristics have not been tested for their effectiveness with children in this study. However, when 80% or more of the websites apply a particular design characteristic, we will speak of a strong convention that seems to be generally accepted by web designers. These strong conventions may well be classified as standards with additional research in which their effectiveness with children has been tested.

In contrast to Nielsen (2004), we will not speak of a convention when only half of the websites presented a particular design characteristic. We set the minimum to speak of a conventional characteristic at 60%, because the characteristic would then be represented by a reasonable majority in the corpus.

3.3.6 Data analysis

After identifying the current conventions used in informational websites for children, the next question that we wanted to answer was whether we could identify relevant subgroups of types of children's websites in the corpus. We did not have clear hypotheses about which types of children's websites we expected to identify in the corpus. Therefore, instead of using the data to test hypotheses on a theory in advance, we made use of Strauss and Corbin's 'Grounded Theory' (Strauss & Corbin, 1990). This is a method in which the theory is inductively developed from the data. The method consists of several stages in which the researcher works from specific 'codes', through 'concepts' and 'categories' to a general 'theory'.

The interpretative process was started by collecting the data set by filling in the coding scheme described earlier in this section (i.e. 'codes') and conducting a hierarchical cluster analysis on the collected data to explore which cases (i.e. websites) formed groups because of similarity on a set of variables. For this cluster analysis, Ward's method was used with a binary Squared Euclidian distance measure because the data set consists of binary codes (i.e. ones and zeros). Websites that were coded the same for a large number of characteristics in the coding scheme formed clusters of websites in the output of the hierarchical cluster analysis (see Appendix B).

3.4 Results part 1: conventions in children's informational websites from the total corpus

In the first part of this section, results will be described for the total corpus. In the second part, results will be presented for the different groups of websites that could be identified through hierarchical cluster analysis.

According to our definition of convention levels, we could identify ten moderate design conventions (60%-80% occurrence) and five strong design conventions (occurrence of 80% or more) in the corpus. In Figure 3.1, the website Kids.gov is presented in which all design conventions that will be discussed in this section are present.

3.4.1 Strong design conventions

First, the strong conventions for children's website design will be presented in Table 3.3 after which they will be discussed in this section.

Table 3.3: Strong design conventions

Strong design conventions for the total of 100 websites	Percentage
1. Main menu available as the primary navigation tool	98%
2. Short texts	93%
3. High contrast between text and background	91%
4. Mouse-over markings of hyperlink click-ability	84%
5. Normal font type and font size	82%

Main menu as the primary navigation tool

Almost all websites provided a main menu on the home page as the primary navigation tool that consisted of a set of horizontally or vertically aligned hyperlinks, both through texts with or without icons (98%). There was no conventional location for the main menu; about one third of the main menus was presented horizontally at the top, one third as a navigation rail at the left and one third through categories in the middle.

Presentation of short texts

Most websites in the corpus presented short texts on the home pages and following navigation pages, with a maximum of approximately 50 words per text block (93%). These high percentages of short texts correspond with Nielsen's guideline to "minimize

Figure 3.1: Example of a website that represents all conventional web design characteristics for children's informational websites mentioned in Table 3.3 and Table 3.4

the amount of text on screens" (Nielsen & Gilutz, 2002). Nielsen says: "Long pages of text frightened users away, even if they were interested in the content."

High contrast between text and background

For most home pages in the corpus the contrast between text and background was high (91%). This standard of high contrast corresponds to Nielsen's guideline to "place text on solid background, because users' reading slowed down substantially when images were overlaid with text, whether tiled or one-picture backgrounds were used" (Nielsen & Gilutz, 2002).

Mouse-over markings of hyperlink click-ability

Most websites presented hyperlinks that were marked as clickable when hovered over by a mouse (84%), for example, by appearance of underlining, altering of the color, or movement (animation) of the link label. This convention corresponds to Nielsen's guideline to 'make clickable items look clickable' and 'add simple visual rollovers to images that can be clicked' (Nielsen & Gilutz, 2002).

Normal font type and size

Most websites (82%) presented a 'normal', highly readable font, such as Arial or Verdana and a font size of approximately 10-12 points. These results correspond to Nielsen's guideline that recommends for children's website design to use simple, relatively large fonts and to avoid using animated text.

3.4.2 Moderate design conventions

In Table 3.4, the ten website characteristics are summed up in order of frequency that are 'moderate conventional', which means that they occured on 60%-80% of the websites in the corpus.

Table 3.4 Children's design conventions

Design conventions for the total of 100 websites	Percentage
1. Service links located horizontally at the bottom of the page	77%
2. Search engine available	72%
3. Ordered page layout (no clutter)	70%
4. Menu presented with text labels (no icons)	70%
5. Solid background color	69%
6. Logo located at the top left corner of the page	67%
7. Separate home button available (apart from logo)	67%
8. Hyperlinks clearly marked as clickable (without mouse-over)	67%
9. Logo functions as a home button	64%
10. Substantial use of graphics	63%

Service links located horizontally at the bottom of the page

On most home pages the service links such as 'contact', 'help' and 'disclaimer', were presented horizontally at the bottom of the page (77%).

Search engine availability

A search engine was available on 72 of the 100 websites, which made the availability of a search engine moderate conventional. This is in contrast to the recommendation of Meloncon *et al.* (2010) not to include a search option on a website for children, because "children have not yet fully developed the intellectual ability necessary to generate relevant search terms (Druin et al., 2009). In addition, giving children the option to search would undermine the process of having them read through the information and explore the website (Bilal, 2001)". However, because Meloncon *et al.* (2010) did not include a search engine on their research website, their study could not confirm or weaken these problems they mention with a search engine for children.

The conventions of the search engine characteristics were computed separately (over a total of 72 websites) and summed up in Table 3.5. Only moderate design conventions were found for search engine design. None of these conventions concern design characteristics of search results (as mentioned in Appendix A, subcategory 2g), such as whether the search results include summaries or snippets of the results pages, whether the URL is presented, whether the result is presented with a relevant image or icon, etcetera. Apparently, there is no strong idea about how search results can be presented best for children.

The presentation of a summary of the search results' content was not conventional. Also, the presentation of a picture in front of each search result was not conventional in the corpus. Furthermore, query formulation support was very rare in the corpus, with only 21% of the search engines that provide spelling suggestions when a query was spelled incorrectly (e.g. Google's 'Did you mean').

We think that these results are surprising, because a summary might support children better than the presentation of a snippet of the content. Also, a picture might support children's use of search results, because of children's preference for images. Absence of spelling suggestions is remarkable, because spelling causes the most problems for children's key word searching (Borgman et al., 1995; Druin et al., 2009).

Table 3.5: Children's search engine conventions

Conventions for the total of 72 search engines	Percentage
1. Results pages offer opportunity of paging search results	72%
2. Results pages present a maximum of 10 results per screen	69%
3. Search engines offer internal search results (instead of external search results from other websites)	68%
4. Search engines accept natural language queries	63%

Ordered page layout

Of the websites in the corpus, 70% were evaluated as ordered and not as 'cluttered'. Clutter is when excess items, on a page, such as a lot of graphics, objects and animations, lead to a degradation of performance when trying to find certain information (Koyani et al., 2006). General guidelines for page design recommend avoiding cluttered pages, because it decreases usability of the page (Lazar, 2006). Younger children especially have difficulties with ignoring irrelevant features. When a web page is cluttered and filled with irrelevant visual features, children will have more difficulties to fixate on relevant items (Brown & De Loache, 1978 about 'visual scanning with children').

Menus presented with text labels

Most websites presented menus with sole text labels (70%). Only a subset of websites presented menus with text labels that were accompanied by icons or menus with icons as hyperlinks without text labels.

Solid background color

Most home pages in the corpus (69%) presented a solid background, which means that the background consisted of one plain color. These findings correspond to Nielsen's guideline to place text on a solid background (Nielsen & Gilutz, 2002).

Logo located at the top left corner of the page

Research on design standards for company websites found that 90 to 100 percent of the pages place the organization's logo in the top left of the page (Adkisson, 2002; Jones & DeGrow, 2011). Also, 84% of the websites in Nielsen's study (Nielsen & Tahir, 2002) placed the logo at the top left, which makes it a general web design standard according to Nielsen. Surprisingly, this location is not a strong convention for children's websites, because only on 67% of the home pages in the corpus could the logo be found in the top left corner. The second occurring location was in the top center of the page (20%).

Separate home button available

The availability of a separate home button was found to be moderate conventional for children's websites (with 67% of the websites). This is an important design characteristic for children, because they tend to go back to the home page when they start a new search task or when they lose track on their current search task.

Hyperlinks clearly marked as clickable

Marking hyperlinks as clickable (without having to mouse-over) was moderate conventional in the corpus, because two third of the websites (67%) presented clear hyperlink

markings. Markings of click ability are underlining, presentation of a button or presentation of a different colored shape.

Logo functions as a home button

Of the websites in the corpus, 64% presented a logo that also functioned as a home button. It is a general guideline for adults that the logo should be a link to the home page (Nielsen and Tahir, 2002). Only five websites in the corpus did not present a home button at all on the website.

Substantial use of graphics

About two third of the home pages in the corpus (63%) presented a substantial number of images. One third of the home pages in the corpus presented only small numbers of images or no images at all. A remarkable finding was that the content pages contained fewer images than the navigation and home pages and that the home pages included most images. This seems to indicate that designers wanted the home pages to be most attractive for children. This substantial use of graphics does not correspond to the use of graphics on websites in general, because both Nielsen and Tahir (2002) and Jones and Degrow (2011) found less use of pictures in their corpora of websites aimed at adults. In other words, it seems that designers for children's websites tend to use more images than designers of websites for adults.

3.4.3 Conclusion about design conventions for children

Many design conventions for children's websites found in this corpus study correspond with guidelines and conventions for general web design. This is not very surprising, because these conventions are often known as general guidelines for web design and because so few guidelines exist for children's websites. Therefore, it is likely that developers of children's websites often follow general design guidelines when designing websites for children. This does not mean that these conventions should automatically be seen as guidelines for children's web design, because more research is needed to test whether these design conventions are suitable for children's search behaviour on informational websites.

However, these identified conventions only represent a part of the analyzed characteristics in the corpus analysis. In the next section, a closer look will be taken at the data from the corpus, to discover whether there were systematic differences for different groups of children's websites. These groups may partly share design characteristics with the total corpus, but may also share characteristics that are specific for a subset of children's websites.

3.5 Results part 2: Categorizing the corpus in three types of children's websites

In this section, the results of the hierarchical cluster analysis will be discussed to identify different categories of children's informational websites. The output of the hierarchical cluster analysis is presented in a dendrogram in which we have set a minimum of ten websites per group (see the 'group identifier' in Appendix B). In that way, five groups of websites were identified from the data (i.e. 'concepts'). From these five groups, three groups clearly represented different categories of children's informational websites. Two groups did not clearly represent a category and were therefore left out of consideration. There were no singletons in the data set, which means that there were no websites that did not belong to any of the five groups.

The three groups that clearly represented different categories of children's informational websites were named: 1) Classic websites, 2) Classical play websites and 3) Image map websites. In this section, we will describe the specific design characteristics of these three groups of websites and we will explain why we have suggested the three mentioned names for the groups. The description of these categories may lead the way to a general 'theory' about the design of children's informational websites.

3.5.1 Classic web design for children

The largest group of websites in the corpus (37 websites, group A in Appendix B) followed guidelines for basic web design, termed 'Classic websites'. According to the literature, the concept of 'Classic web design' aims at simplicity, consistency and focus in web design. The layout of the pages on such websites is kept minimal; key elements of the pages are the center of attention and different pages have the same layout as recommended by Brinck *et al.* (2002). Page components are located on conventional locations as recommended in the literature on usability guidelines for web design (Koyani et al., 2006; Nielsen & Tahir, 2002).

Many of the Classic websites in the corpus presented a directory of main categories through which the user could browse to relevant information. This can be seen, for example, on the home page of Dibdabdoo.com (see Figure 3.2), which is a good example of a typical Classic website. In Table 3.6, the strong and moderate conventional characteristics are presented for the Classic websites. Characteristics that are conventional for all websites in the corpus are not presented, because they were already discussed in the previous section.

Figure 3.2: Example of a Classic website for children from Dibdabdoo.com

Table 3.6: Conventions for children's Classic web design (N=37)

Strong conventions	Percentage
Ordered pages (no clutter)	97%
Solid backgrounds	92%
Consistency in menu structure	92%
Pages with low screen density	89%
Link labels only presented with text	86%
Consistency in layout	84%
Moderate conventions	
Search engine availability	76%
Main menu on a conventional location (at top or left of page)	73%
Low use of graphics	73%
Low use of animations	73%
Clear (non-ambiguous) main categories	68%

Clean and ordered design

The Classic websites in our corpus were characterized by a solid background color (mostly white), little clutter on the pages and low screen density. 'Density' is the percentage of the screen filled with text and graphics. Websites with a lot of 'empty space' have a low screen density. Children will not get lost easily on the Classic websites, because of the consistent menu structure and layout throughout the websites. Common elements, such as the logo or the main menu are placed at a conventional location on most of the Classic websites. This clean and ordered page design might support children to coordinate and control their scanning activities, because there are no irrelevant distracting visual features.

Low use of multi media

The Classic websites contained small numbers of images and animations. The same counts for the use of video and audio, which was also very limited on the Classic websites. The perceived hedonic quality of these websites might therefore be low. Children might not be stimulated by the websites because of lack of originality, creativity and excitement of visual features such as pictures and animations. Therefore, the websites might be experienced as 'boring' by children, because of the small number of pictures and illustrations.

3.5.2 Classical play for children

A second category of children's websites that was discovered in the data was termed 'Classical play'. With these websites a classic design approach was combined with a playful, visual design approach. This was already recommended by Harbeck and Sherman (1999): children's websites should be simple, clean and concrete, yet also active

and enjoyable. A 'playful approach' means that more effort is spent on the design of graphics, color, and games (Meloncon et al., 2010).

These 'Classical play websites' (14 websites in the corpus; group B in Appendix B) shared characteristics with both the Classic websites and the Image map websites. They were characterized by a traditional arrangement of information, just like the Classic websites, but shared a playful approach with the Image map websites, which will be described in the next section. In Table 3.7, the conventional characteristics are presented for the Classical play websites.

A good example of a 'Classical play' website is BAM.gov (see Figure 3.3), a website of the Department of Health and Human Services of the US Government, especially designed for children, with a substantial number of graphics, animations and colors.

Table 3.7: Conventions for Classical play web design (N=14)

Strong conventions	Percentage	Moderate conventions	Percentage
High number of graphics	100%	Main menu on a conventional location (at top or left of page)	79%
Animations presented	100%	Search box at top right of page	79%
Link labels presented with text and icons	100%	Pages with high screen density	64%
Clear main categories	100%	Playful fonts	50%
Focal point available	93%		
Logo as home button	93%		
Consistency in menu structure	93%		
Consistency in layout	86%		

Use of multimedia

All Classical play websites presented a substantial number of images. Large *et al.* (2002) report that children responded positively towards pictures and that children were critical about websites that did not present pictures. However, they also say that children appreciated pictures, when they added value to the content, for example, to illustrate something.

All home pages in this group included animations. The use of animations is an important contribution to the playful character of the websites, because motion is an extraordinarily compelling visual quality that attracts the user's attention (Williams, 2000). However, children might also experience difficulties with scanning and parsing web pages that are presented with a lot of visual features. Fixating on relevant items on such web pages becomes more difficult for children when the screen is filled with distracting visual features.

Also, some websites in this group make use of audio (6 of 14 websites). Some websites present sounds when users hover over hyperlinks. Nielsen and Gilutz (2002) report that children appreciated that kind of sounds: "Adding audio rollovers provides an experience enhancement that kids enjoy, especially if the sounds are funny."

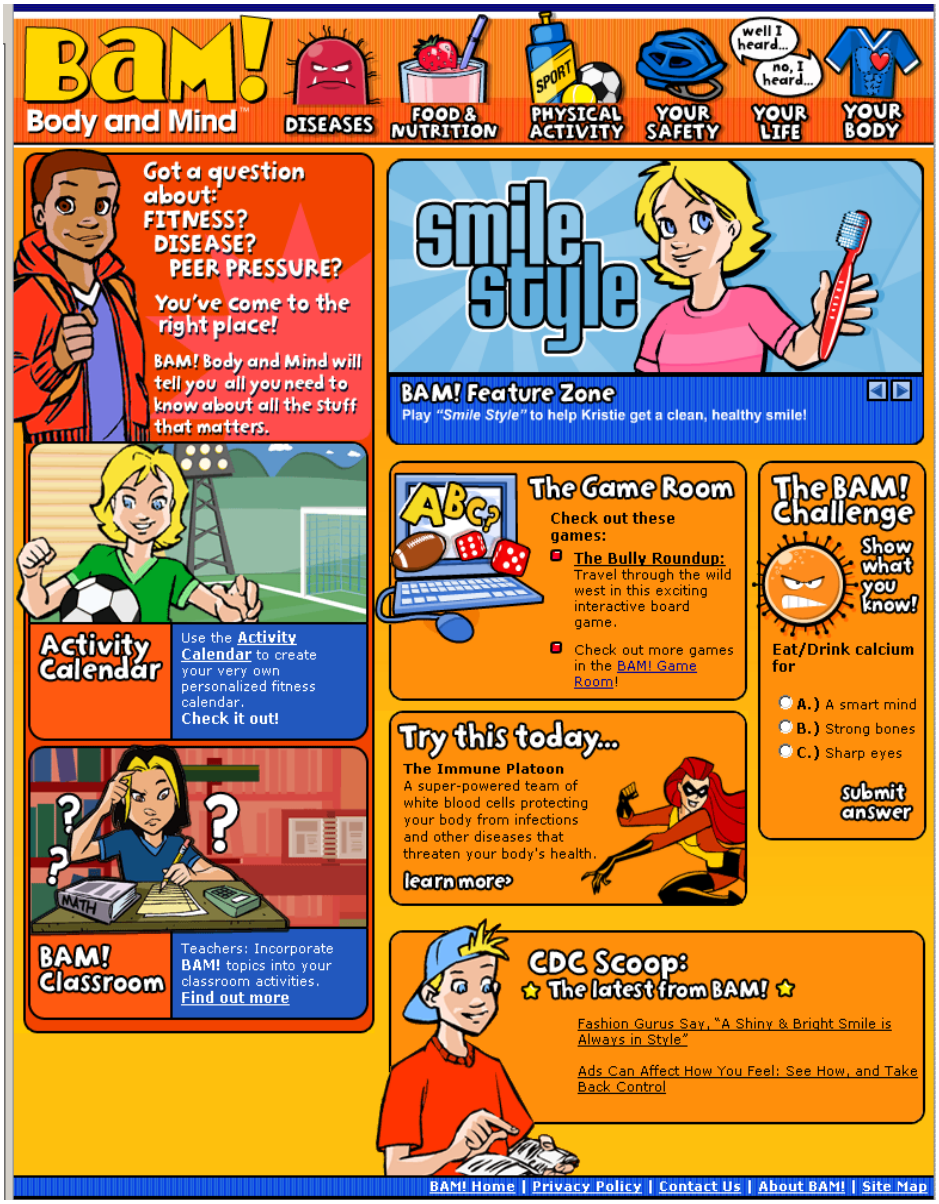


Figure 3.3: Example of a Classical play website for children from Bam.gov

Icons as menu labels

All websites in the Classical play group presented menu labels accompanied by icons. Most of these websites (93%) presented icons that were meaningful, which means that it was clear right away to the assistants which category the icon stood for.

Consistency and main categories

Menu structures and layout on the Classical play websites were consistent throughout the pages, which supports children's navigation and orientation. Also, the main categories on all websites were scored as concrete, which means that the assistants had the right expectation of the information that could be found after selecting the categories. As discussed in the section about children's cognitive development, children have difficulties with browsing through abstract categories. Because of the concrete main categories, difficulties with browsing categories should have been limited with the Classical play websites.

Conventional location of common elements

Many items on the Classical play websites were presented at conventional locations. The logo was usually presented in the top left corner of the page, which was also a home button for almost all websites in this group. The main menu was presented horizontally at the top of the page (as can be seen on the home page in Figure 3.3). The search box was usually presented in the top right corner of the page and the service links horizontally at the bottom of the page as a navigation footer. These conventional locations of items can also be seen on the home page in Figure 3.3. Another common characteristic of this group of websites was the presentation of a clear focal point on the page, often in the form of an image or animation, which immediately drew the user's attention. In Figure 3.3, the picture with 'smile style' is an example of such a focal point.

3.5.3 Image map websites for children

In the third group, which consisted of 15 websites in the corpus (group C in Appendix B), no classic web design characteristics were used. The visual design and navigation of the websites in this group were based on Image maps that incorporated objects or locations that children know from real life or from fiction. The child could explore a tableau of real life or fictional objects in search of specific information. This makes information search a playful experience. Meloncon *et al.* (2010) also made use of Image maps for navigation in their educational website for children.

The Image maps might support children in using a system that is new to them. They can relate to an environment that they already know, which makes it more intuitive for the child to navigate and manipulate within this space (Barfield, 2004). Examples of such Image maps are a concert hall (The New York Philharmonic Kidzone), a tropical island (BR-Kinderinsel), a school bus (The Magic School Bus) or a space shuttle with planets (Kids Health Galaxy.com; see Figure 3.4). In Table 3.8, the conventional characteristics are presented for the Image map websites.

Table 3.8: Conventions for Image map web design for children (N=15)

Strong conventions	**Percentage**
High numbers of graphics	100%
Use of animations	93%
Non-solid backgrounds	80%
Separate home button	80%
Image map navigation	80%
Consistency in layout	80%
Moderate conventions	
Use of audio	60%
Consistency in menu structure	60%

Use of multimedia

The amount of multimedia use was high on these websites. Almost all of these websites presented a screen-filling graphic (i.e. Image map) and made use of animations. Also, many of these websites made use of audio, such as the spoken introduction by the character 'Joey' in Figure 3.4.

Page layout

On most of the websites in this group, the logo was not a home button. Instead of that, a separate home button was often provided. The main menu and the search box were mostly included in the visual metaphor on a non-traditional location. In 80% of the cases, the background on these websites was non-solid, because of the visual metaphor that covered the entire screen. Many of these websites also made use of playful fonts.

Icons as category labels

Some Image map websites presented icons as category labels without textual labels. Textual labels appeared with these items, when the user hovered over them with the mouse (i.e. 'mine sweeping'). Nielsen found that children liked to mine sweep pages in search for hyperlinks (Nielsen & Gulitz, 2002).

Search engine availability

Remarkably, the number of websites with a search engine was lowest on the Image map websites. An explanation might be that these websites were especially aimed at 'exploring' the website in search of the right information. As discussed in the section about children's cognitive development, the emphasis on browsing categories instead of keyword search might have a positive effect on children's search performance, because of the lower cognitive load of browsing categories compared to keyword searching.

Figure 3.4: Example of an Image map website on Kids Health Galaxy.com

3.5.4 Groups of websites left out of consideration

Two of the groups identified from the hierarchical cluster analysis are not addressed as specific categories of children's informational websites and left out of consideration in this article. The websites in these two groups shared a design approach that did not clearly point at a clear category. The first of these two groups can be characterized by cluttered pages, high density and a lot of textual elements. It seems as if the designers of these websites tried to put as much information on one page as possible. Regarding the graphical character of these websites, they related to the Classical play websites, but they were too crowded for this group of websites. The second group that was left out of consideration can be characterized as poorly organized, with little consistency in menu structure and layout and scarce use of orientation cues. Thus, we do not claim that all children's websites can be categorized in the three main design categories that will be described in this section.

3.5.5 Conclusion: three types of children's websites

In this section, three categories of children's websites were identified in the corpus of children's informational websites, which differ in the amount of playfulness added to the web pages. No playful features were added to the Classic websites. The Classical play websites were only playful in visual appearance. In the Image map, the total interaction design was playful. On these websites, there were not only visual playful features, but also functional playful features (e.g. navigation tools).

3.6 Conclusion and discussion

The results of this corpus study indicate that most design conventions that were identified in the corpus of children's informational websites correspond to general web design guidelines. The answer to *RQ 1: Which design conventions on interfaces for children are specific for children and which of them are general design conventions?* is that design conventions for children's search interface design are mostly general design conventions and not specific for children. The conventions identified in this chapter were not tested and validated as standards for children's web design in this study. However, most of the identified conventions correspond to the literature about how children interact with websites (Nielsen & Gilutz, 2002; Meloncon et al., 2010). Design conventions such as use of short texts, high contrast between text and background, marking of hyperlink click ability, conventional location of items as the logo (top left of page) and service links (bottom of page), ordered page layout, a solid background and a consistent layout through the website are not only proved helpful for adults, but also for children.

Still, many websites in our corpus had design characteristics that did not belong to a set of generally accepted design conventions. Apparently, designers do not often agree about what kind of design characteristics are suited for children's search behaviour. For example, 58% of the websites required scrolling up and down the page to view all information available. Meloncon *et al.* (2010) do not provide conclusive proof in their study about children's preferences regarding page length and their willingness to scroll, while Nielsen and Gilutz (2002) recommend no scrolling on web pages for children. Further, the labels of the main categories were analyzed as 'vague or ambiguous' on 42% of the websites. Also, a third of the websites used a non-solid background, which is a risk for a children's website, because reading and scanning slows down substantially when the background is filled with images (Nielsen & Gilutz, 2002).

Furthermore, almost a third of the websites (30%) was analyzed as being 'cluttered' and on 41% of the websites the pages were characterized as having a high density. Children will have more difficulties to fixate on relevant items on these cluttered pages,

because the younger the child, the more attention he is likely to give to stimuli that are irrelevant to the original task (Brown & De Loache, 1978).

Most of the websites in the corpus did not offer orientation cues, such as marking of the current location in the menu (only 37% of the websites), marking the color of the page as orientation cue (23% of the websites), or by offering a navigation path or breadcrumb trail (23% of the websites). This could also cause problems for children while searching for information on a website, because children's meta-cognitive control is still in development. Therefore, they may have problems in monitoring their location in the website.

More research is needed to validate the design conventions as guidelines for children's informational websites and to find out in what way design characteristics that are not clearly conventional from this study, are suited for children's search behaviour.

3.6.1 Types of children's informational websites

Lavie and Tractinsky (2004) examined user's perceptions of website aesthetics and found two main dimensions of aesthetics: classical aesthetics and expressive aesthetics. They found that classical aesthetics are characterized by clear and orderly design and are closely related to general design rules. This characterization of classical aesthetics corresponds to the category of Classic websites in our corpus, which were classic in both visual design and in navigation structure. By contrast, expressive aesthetics can be characterized by creativity and originality and by the ability to break design conventions, which corresponds to the more playful website categories. The Classical play websites were expressive in visual design through presentation of many colors, images and animations, but the navigation structure of these websites followed classical aesthetics. The Image map websites were expressive in both visual design and in navigation structure. With a playful navigation structure, for example, there was no textual menu, but the user had to 'mine sweep' through visual elements in search of a relevant category.

Other research reports that interactive products are perceived and evaluated along two different dimensions: pragmatic and hedonic quality (Hassenzahl, Schöbel & Trautmann, 2008; Van Schaik & Ling, 2008), both from a subjective user-centered perspective of quality of use. Pragmatic quality refers to the perceived functional abilities of the website to achieve do-goals, such as 'finding information for a school assignment'. In contrast, hedonic quality refers to the perceived pleasure-producing qualities of the website to achieve be-goals, such as 'being special', or 'being related to a particular group'. Two forms of hedonic quality are stimulation (e.g. the product is original, creative, innovative or exiting) and identification (e.g. the product is integrating, professional, inclusive or presentable). The variation in playfulness of children's websites might also have

implications for this difference in children's perception of pragmatic and hedonic quality of these different categories of children's web design.

Classic websites

Based on the theories of Lavie and Tractinsky (2004) and Hassenzahl *et al.* (2008), we would expect the Classic websites to be perceived as having high usability (i.e. pragmatic quality) and to be supportive of children's search behaviour because of the clear and ordered layout and structure. Cognitive developmental psychologists (Brown & Deloache, 1978) reported that the younger the child, the more attention it is likely to give to stimuli that are irrelevant to the original task. Irrelevant stimuli that could cause problems for children while scanning a web page or searching for relevant information are absent on Classic websites. Therefore, perceived pragmatic quality and usability of these websites might be high.

The perceived hedonic quality of the Classic websites might be low because of the low number of playful characteristics, such as graphics, animations, audio, colors, games or playful navigation forms. Children's motivation to search information on these websites might be low because of a lack of stimulation through creative and exiting elements.

Classical play websites

The Classical play websites that combine the classical and playful approach might be the best solution for children; perceived usability might be high because of the clear and ordered layout and structure and perceived hedonic quality might be high because of the playful features. However, it may also be the other way around; the playful features might be distractive for children, which decreases perceived pragmatic quality of the websites. And the playful elements might not be appreciated enough by children to have a positive effect on perceived hedonic quality.

Image map websites

The implications for the Image map might be just the opposite of the Classic websites. The perceived usability might be decreased by the use of Image map websites because common elements are not located on conventional locations and a child might not recognize how to navigate through the environment if he is used to more basic web design. Also, scanning and searching might be problematic, because the playful interface offers distracting stimuli that are irrelevant to the search task. The perceived hedonic quality of Image map websites might be high, because of their creative, innovative and playful character. Children's need for novelty and change and relatedness (Hassenzahl et al., 2008) might be fulfilled by these non-conventional forms of navigation. These Image maps can be seen as metaphoric sites: they 'transport' the children to the world of the website, that is, the school bus, boat or airplane. This allows the child to identify totally

with the interface, which might also cause perceived hedonic quality to increase. An interesting question is whether this intended effect of transfer indeed takes place: do children actually step into this new world, and does this have a positive effect on their evaluation of the website?

3.6.2 Future research

By varying the amount of playfulness on the dimensions of visual design and navigation structure, something special is happening in the field of children's information and interaction design of which we do not know the implications for children's cognitive and affective states. For a long time, research on children's information design was limited to the usability of the environments and limited to pragmatic questions such as how to design interfaces that are effective and efficient for children's search behaviour. However, by adding playful elements to the design, such as images, games and animations, we see that designers are trying to influence user's emotions and affective states. This development can also be seen with the informational genre for adults (e.g. company websites) in which designers try to influence the user's affective states.

This change from an objective, pragmatic to a subjective, hedonic perspective on information design is in line with the concept of 'user experience' (UX), which has emerged in the field of human-computer interaction (HCI) and interaction design (Hassenzahl and Tractinsky, 2006). User experience is associated with meanings ranging from traditional usability to beauty, hedonic, affective or experiential aspects of technology use. It is about the maturation of technology in which interactive products not only become more useful and usable but also more fashionable and fascinating things to desire. The user experience focus is on aspects beyond the functionality of design. It is about designing for quality experiences and pleasure rather than merely preventing usability problems, as was the case with interactive products in traditional HCI (Hassenzahl & Tractinsky, 2006).

We are optimistic that the coming years will mark a turning point in children's digital information design. It seems likely that more research will be conducted with more advanced technologies and with many different digital devices that study the relation between affective and cognitive factors during information processing.

In Chapter 4, we will report the results of an explorative follow-up study on children's interaction with the three identified types of children's informational websites, represented by three existing websites. In Chapter 5, the results of an experiment will be reported, in which we studied both cognitive and affective states of children using exactly the same web content, but manipulated in three different design versions, following the three categories that we discovered in this corpus study. These three design version were: a Classic version, a Classical play version and an Image map version. This study gave us more insight into the quality of children's informational websites

from both a performance perspective and an evaluation perspective and the interplay between these two perspectives during children's search performance.

ZOEK

Chapter 4

The design of informational interfaces for children: exploring effects on performance and affective experience

4.1 Introduction

Digital media are of growing importance for children's education and their information searching. Information seeking has moved from traditional media, such as school books and library books, to digital media, such as online encyclopedias and informational websites (i.e. information *search*). Most research on children's online behaviour is aimed at children's safety and on how to protect them from harmful information that is available on the World Wide Web (Wishart, 2004; Valcke, De Wever, Van Keer & Schellens, 2011). Less attention is given to the issue of how to support children in effective information searching when they search for information in digital environments and the difficulties that children experience while searching for information (Borgman et al., 1995, Schacter et al., 1998).

Furthermore, to our knowledge, the aspect of visual design of digital environments has not yet been addressed in previous research on children's information searching. In this chapter, we will present a study on the effects of different types of informational Web design for children on children's search behaviour and performance, on their emotional reactions and on their evaluation of the search interfaces.

In the corpus study reported in Chapter 3, in which we analyzed 100 informational websites for children, we found that there are three design categories for this website genre. The first type is a classic type of Web design that is typically characterized by classical aesthetics (Lavie & Tractinsky, 2004), such as clear and orderly design (for example, see Figure 4.3). We named the second type Classical play Web design; classical because of the clear and ordered design of the navigation structure, but playful because of expressive visual aesthetics, such as the presentation of many colors, images and animations (for example, see Figure 4.4). The third type of Web design was named Image map design. This term was also used by Meloncon, Haynes, Varelmann and Groh (2010). Both visual design and navigation structure in this type of Web design were characterized by expressive aesthetics. There is no textual menu, but the user has to 'mine sweep' through an Image map in search of a relevant category of information (see Section 4.4.2 for an example of an Image map website).

We assume that Web designers have good reasons for using all three types of Web design: aiming at children's positive perception of usability, positive evaluation and positive perception of hedonic quality of the interface. The study reported in this chapter is a first exploration of the actual effects of these different types of web design on both usability and fun (that is, pragmatic and non-pragmatic aspects of user experience) and on the suitability of different methods to measure usability and fun with children. In the next section, we will describe the interplay between different components of user experience, such as usability and fun (Thüring & Mahlke, 2007) in more detail.

The children in this explorative study worked with five websites in total. The children's interaction with three of these websites was analyzed in depth and the results of this analysis will be reported in this chapter. These three websites represent the three design categories for this website genre, as identified by the results of the corpus study: a classic website (schoolbieb.nl), a Classical play website (willemwever.nl) and an Image map website (kids.kennisnet.nl). The fourth website that was added to the study was Google, which functioned as a baseline for our study. We wanted to be able to compare children's search behaviour on websites that were especially designed for children to their search behaviour on a website which they are most experienced with (Pijpers, Marteijn, Bosman, Van den Berg, Dijkerman, 2008), but that was initially designed for adults (i.e. Google). The fifth website (wikikids.nl) was added as a 'filler' stimulus in this study. This website belonged to the group of websites that did not clearly represent a category in the corpus study (see group D, appendix B). Children's search performance on this website is further left out of consideration in this explorative study. We will argue at the end of this chapter whether this website should be represented in the experiment that will be reported in Chapter 5.

In this chapter, we will focus on children's search performance on and evaluation of the three websites that represent the three categories identified by the corpus study reported in Chapter 3. Before we report on the design, methods used and results of the explorative study, we will first discuss prior research that is relevant for this study.

4.2 Theoretical background

4.2.1 Focus on pragmatic quality and cognitive efficiency

The literature on children's interaction with digital products over the last decade is dominated by pragmatic issues, such as the utility and usability of technical systems (Thüring & Mahlke, 2007). Mostly, performance-based methods were chosen to assess two usability components: effectiveness and efficiency (Bilal, 2002). A third component of usability that was often studied is user satisfaction. However, these subjective judgments of the use of the system are also mostly based on the efficiency and effectiveness of system usage (Lindgaard & Dudek, 2003).

In studies on the cognitive efficiency of websites for children, researchers have often found that children experience difficulties while using both searching and browsing tools in digital environments. These tools do not take into account children's cognitive and motor skills. Many websites for children offer keyword-based search engines as an important search tool. However, formulating a search query is difficult for children, because they have little knowledge to 'recall' concepts or terms from their long-term memory (Borgman et al., 1995; Hutchinson, Druin, Bederson, Reuter, Rose & Weeks,

2005). Besides, for searching relevant documents using keyword search, correct spelling, spacing and punctuation are needed, because most search engines do not offer spelling corrections or suggestions. Children often make spelling errors (Borgman et al., 1995). Deciding on a single keyword is also difficult for a child, because children tend to use a full natural language query. Unfortunately, most search engines on children's websites do not accept natural language queries (as reported in Chapter 3).

With browsing, children first of all have trouble finding the right category, because they have little domain knowledge to decide which category is most relevant. In addition, problems with browsing tools are mostly the result of a lack of vocabulary knowledge.

Children often have difficulties understanding abstract, top-level headings, because their vocabulary knowledge is not yet sufficient to understand such terms (Hutchinson, Bederson & Druin, 2006). Unfortunately, most websites for children offer quite abstract or vague top-level headings, because so many subjects have to be covered in these main menus. For example, some of the top-level headings of KidsClick.org that are presented in Figure 4.1, are quite abstract and ambiguous, such as "reading, writing & speaking" or "weird & mysterious". Children are able to use hierarchies to locate information. However, they may experience difficulty in conceptualizing abstract concepts and traversing deep multilevel hierarchical structures. The deeper the hierarchies, the more likely children will get lost (Bilal & Wang, 2005).

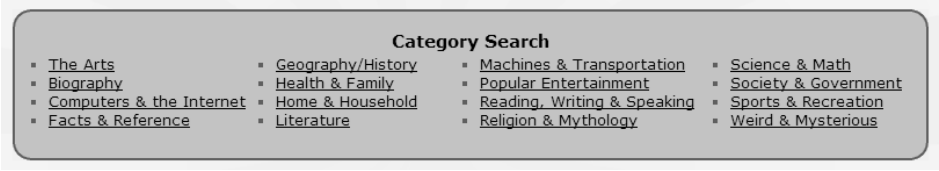


Figure 4.1: Top-level headings of KidsClick.org

4.2.2 Model for web navigation

To examine children's digital information searching, a model was needed to divide the search process into several constituent processes. The Comprehension-based Linked Model of Deliberate Search (CoLiDeS) (Kitajima, Blackmon & Polson, 2000) divides the search process in attending a web page, parsing a web page, focusing on an area and selecting a relevant entry. This model assumes that comprehension of texts and images is the core process underlying Web navigation. It is inspired by the concept of 'information scent' (semantic relevance of screen objects to users' goals) and emphasizes the semantic dimension of Web navigation which assumes that the process of relevance assessment is central to web navigation. Information scent is measured based on three factors: semantic similarity, frequency and literal matching. Semantic similarity is calcu-

lated based on co-occurrences between words and documents with the aid of a machine learning technique termed latent semantic analysis (LSA).

4.2.3 Affective user experience

Usability is not all that matters in human-technology interaction. It is also important how children feel during task performance and how they respond emotionally to a system (Thüring & Mahlke, 2007). In 2002, Dillon already proposed that user satisfaction is likely to be influenced by other factors such as personal experience with technology and the aesthetics of system design. The field of user experience (Hassenzahl & Tractinsky, 2006) argues for a broader perspective of user experience than only usability. Besides the perception of instrumental qualities (that is, pragmatic quality / usability), the perception of non-instrumental qualities is also important, such as visual aesthetics or haptic qualities (that is, hedonic quality). Furthermore, it is important to know how people feel while using a system. "Stimulating emotions in the online environment is especially important because communication in the Internet occurs mostly through limited affective bandwidth" (Picard & Klein, 2002). Thüring and Mahlke (2007) present a model that integrates these components of user experience into a framework, which they call the CUE-model, where CUE stands for Components of User Experience (see Figure 2.12). System features (such as functionality and website design), user characteristics (such as knowledge or skills) and the characteristics of the task and context determine the interaction with an interactive product. In the course of this interaction, user experience is gained which comprises of three components: perception of instrumental, pragmatic quality (for example, ease of use, controllability, effectiveness, learnability), and perception of non-instrumental, hedonic quality (for example, visual aesthetics, haptic quality, identification, stimulation; a definition of hedonic quality is given in Section 2.2.4). The perception of both instrumental and non-instrumental qualities is likely to influence the third component of user experience: the emotional reactions to the interaction with a product (subjective feelings, motor expressions, physiological reactions). All three mentioned components have an influence on the overall evaluation of the interactive product.

The emotional side of user experience is still often neglected in the research on child-computer interaction. Especially with children, we expect substantial influence of emotional experiences on their system evaluations during system usage. In a study to gather opinions in child computer interaction, Read and MacFarlane (2006), for example, prove that the major factor in a child's decision about whether they want to use an interactive product again, is how much fun it was instead of ease of use.

4.2.4 Emotional reactions

As mentioned in the CUE-model, one of the three components of user experience is 'emotional reactions'. Following Scherer's (1984) componential theory of emotion, Desmet and Hekkert (2007) apply this theory to what they call 'product experience', with which they refer to an experience that is affective. They report product experience as a multi-faceted phenomenon that involves subjective feelings, behavioural reactions, expressive reactions, and physiological reactions. Where Scherer (2001) mentions appraisal as a cognitive component of emotional response, Desmet and Hekkert (2007) see an emotional response as elicited by appraisal (that is, an evaluation) of a product as potentially beneficial or harmful.

To categorize experienced emotions, a set of standard emotion states is presented, named the Geneva Emotion Wheel (GEW) (Scherer, 2005). The development of the GEW is based on the dimensional emotions approach, first initiated by Wilhelm Wundt (in his 'Outline of psychology', 1896 in Scherer, 2005), who suggested that subjective emotions can be described by their position in a three-dimensional space formed by the dimensions valence (positive-negative), arousal (calm-exited) and tension (tense-relaxed). Many modern dimensional theorists limit themselves to the valence and arousal dimension, because of the difficulty of identifying a third dimension (Russell, 1980). The GEW is also based on this two-dimensional space.

Not all emotions mentioned in the GEW can be experienced in relation to products, such as digital environments that are the subject of this study. For example, pity or compassion, guilt or remorse or envy or jealousy are not affective responses than can be experienced in user-product interaction. Therefore, Desmet (2008) designed a circumplex model of core affect with product relevant emotions (see Figure 4.2), based on Russell's circular two-dimensional model of valence (on the horizontal axis in Figure 4.2) and arousal (on the vertical axis in Figure 4.2), just like the GEW.

4.3 Research questions

We will address the following research questions in this chapter (see A, B and C below) to explore the effects of different types of web design for children on children's performance and emotional reactions. This will help us in answering *RQ 2* and *RQ 3* of this dissertation.

RQ 2. *What is the effect of design characteristics of informational websites for children on their search performance and on their appreciation of the search interface?*

RQ 3. *What is the relation between children's search performance and their affective responses towards the search interface?*

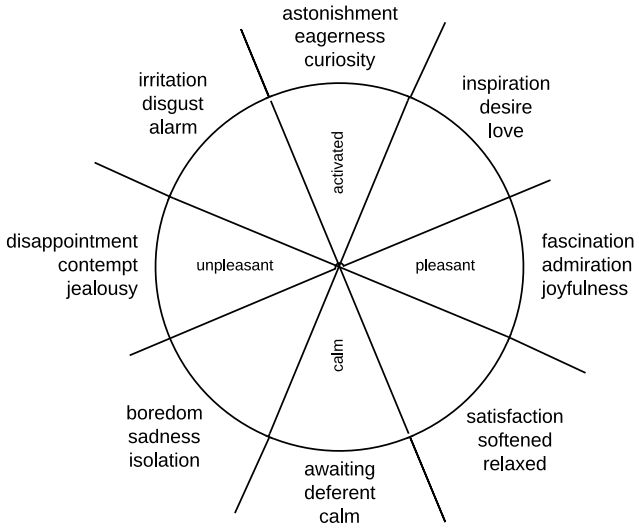


Figure 4.2 Circumplex model of core affect with product relevant emotions (Desmet, 2008)

Research questions in this chapter

RQ - Chapter 4-1. Which of three types of informational websites is most usable for the children?

RQ - Chapter 4-2. What effects do the three different types of website design have on the emotions expressed by the children?

RQ - Chapter 4-3. How do pragmatic quality and experienced emotions relate to each other, and is there an interaction with difference in design characteristics of the websites?

To answer these three research questions, we analyzed the recorded data of children's search performance in this study both from a performance perspective (i.e. pragmatic usability) and from an evaluation perspective (i.e. perceptions of pragmatic and hedonic quality).

4.4 Method

Following the research questions, this study had three concrete goals: 1) to measure the pragmatic quality of these websites and the cognitive efficiency with which children can interact with these systems, 2) to measure the affective side of the user experience by the children by observing their facial and vocal expressions of emotions and 3) to explore the suitability of the methods used for future research. All used methods, procedures and instruments were pilot-tested in the lab at Utrecht University before we

conducted the actual explorative study. The study was not set up as an experiment, but as an explorative study. A within-subject design was used, which means that all participants conducted search tasks on the three websites. The results of this explorative study have been used to develop an experimental follow-up study (see Chapter 5).

4.4.1 Participants

For our study, we approached a primary school in the Dutch province Utrecht, from which we knew that the Internet is an important and frequently used instrument to search information for school assignments in the classroom. We wrote a letter to the parents of 35 children from two classrooms and asked them for their consent for participation of their child. Three parents did not give their consent. In total, 32 children participated in our study; 27 children from a classroom with children from fourth to sixth grade and five children from a classroom with children from first to third grade. In Table 4.1, ages and gender of the children is specified. The study was conducted in March 2010.

Table 4.1: Age and gender of the participants

Age	**Boy**	**Girl**	***Total***
8	0	2	*2*
9	3	5	*8*
10	5	5	*10*
11	3	5	*8*
12	0	4	*4*
Total	*11*	*21*	*32*

4.4.2 Study objects

We conducted the research with children in the Netherlands. Therefore, the websites studied in the research were in Dutch. The websites represented the real information world in basic main categories, such as animals, sports, music, arts, nature, history, etcetera. We selected three websites that met these criteria and that represented one of the three main categories of Web design for children identified in our corpus study (Chapter 3) and described in the Introduction: a Classic website (see Figure 4.3), a Classical play website (see Figure 4.4) and an Image map website (see Figure 4.5). In table 4.2, the website characteristics of the three websites are presented.

4.4.3 Data collection methods

The study was conducted by the author of this dissertation (the test instructor), in a quiet room in the school during school hours. Each child participated individually and the duration of the sessions per child ranged from 30 to 45 minutes. Our data collection methods were both quantitative and qualitative.

Table 4.2: Selected websites and their website characteristics

Type of website	Arrangement of information	Menu structure	Clutter / density	Layout
1. Classic	Traditional Ordered	Taxonomical Textual menu labels Consistent throughout website	Little clutter Ordered pages Low screen density	Little use of graphics and animations / Solid background color
2. Classical play	Playful Build of blocks	Abstract main menu Menu labels presented with text and icons Basic categories on 2nd level Consistent throughout website	A lot of clutter High density	High numbers of graphics and animations Focal point available
3. Image map	Within a playful 'Image map'	Navigation metaphor Game-like navigation design Abstract main categories	Little clutter High density	High numbers of graphics Use of animations Non-solid backgrounds

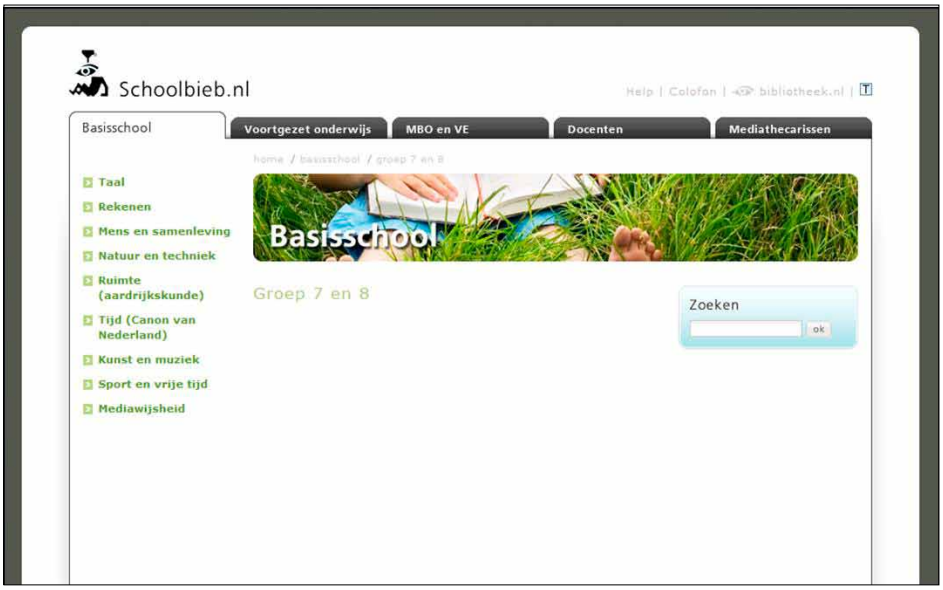


Figure 4.3: Screen display of the Classic website (URL: schoolbieb.nl)

Quantitative data collection

Each child was asked to fill out a questionnaire at the beginning of the session. The questionnaire was a profile survey in which children were asked about demographical data, such as age, grade level and gender. They were asked about their computer experience: frequency of PC use, frequency of using the Internet, activities on the Internet (such as playing games, watching movies, etcetera), and frequency of online

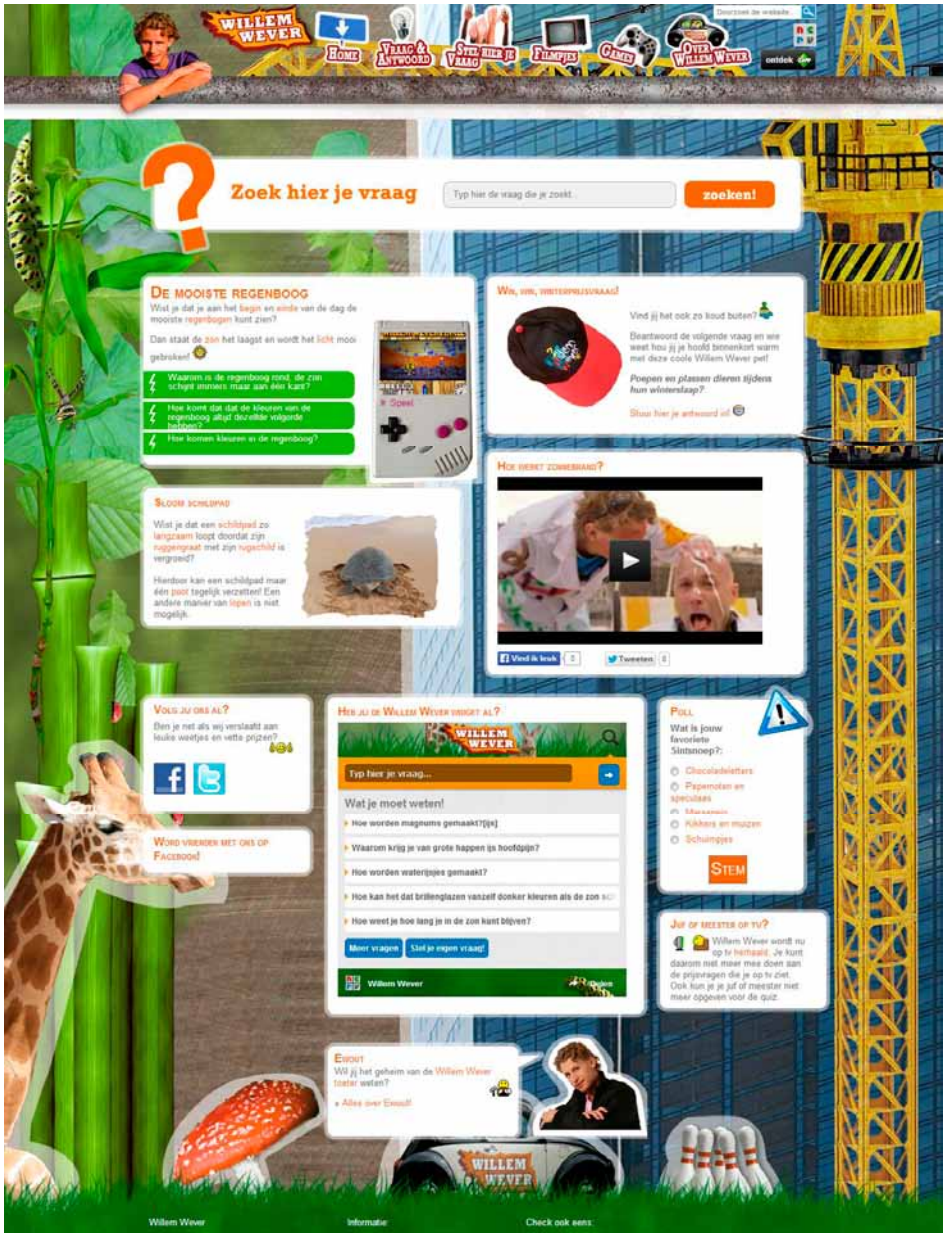


Figure 4.4: Screen display of the Classical play website (URL: willemwever.nl)

information searching. Further, they were asked about their experience with the websites for children that were selected for this study.

Quantitative data during the task performance was logged by recording all browser activities during the search sessions of the children. This method is unobtrusive and can run alongside other evaluation methods (Markopoulos, Read, MacFarlane & Hoysniemi,

Figure 4.5: Screen display of the Image map website (URL: kids.kennisnet.nl)

2008). For each conducted search task, the amount of time and number of clicks that were needed to find the information was measured. From the number of clicks needed, the deviation of the optimum navigation path could also be measured for each search task. For each task, success scores were determined about how successful children were in finding the relevant information on the websites.

After performing each search task, the child was asked to evaluate the difficulty of performing the search task (subjective appraisal scores of instrumental quality) on that particular website on a 'Visual Analogue Scale' (VAS) with evaluations from 'very easy' to 'very difficult' (see Figure 4.6). Our VAS of difficulty of the tasks was inspired by the Smileyometer of Read, MacFarlane and Casey (2002).

At the end of every session, the test instructor asked the child to rank the websites. The website that the child definitely would use the next time for information search had to be ranked as 1, the one that the child would use after that had to be ranked as 2, etcetera. This ranking of activities (or in our case of digital products) is termed comparative fun by Read *et al.* (2002). Our ranking of the overall appraisal of the websites was inspired by their Fun-Sorter.

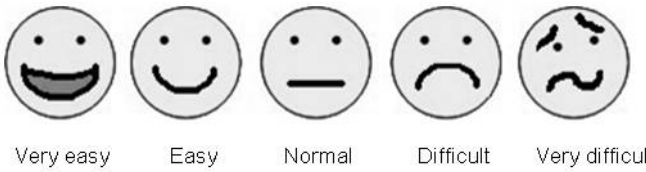


Figure 4.6 Example of the 'Visual Analogue Scale' for expressing experienced difficulty

Qualitative data collection

We collected the qualitative data by using a structured observation method to observe children's performance on the research tasks, during which notes were made of remarkable observations. We did not stimulate children to think aloud during task performance, because we do not think that most children are particularly capable of thinking aloud during their sessions. We also did not want to 'disturb' their natural search behaviour by enlarging the cognitive load of the search tasks by encouraging them to think aloud during information searching (Markopoulos et al., 2008). However, to learn more about children's affective experience, interventions were made during the sessions by asking a neutral question after each search task ("Can you tell me what you found easy / difficult when conducting this task?"), termed post-task interviews (Markopoulos et al., 2008).

During the sessions, all browser activities were recorded and, more importantly, the children's eye movements on the screen during the task performance were recorded using a Tobii 1750 remote eye-tracker (Tobii Technology, Danderyd, Sweden) and the software named Studio. This eye-tracker is a free standing, non-invasive device which resembles an ordinary computer screen. Also, a video display of the child and an audio recording of the spoken comments of both the children and the test instructor were recorded during the sessions.

4.4.4 Procedure and tasks

At the start of every session, the child was asked to sit down in front of the computer screen (i.e. Tobii eyetracker screen). Every step of the procedure was written down in a strict protocol, so that the procedure would be the same for every child. First, the test instructor explained the goal of the research session to the child and the tasks that the child would be asked to conduct on the different websites. After that, the test instructor asked the child to fill out the questionnaire about prior experience with computers and the Internet.

The next step in the session was the calibration of the eye-tracker to locate the position of the child's eyes. After the calibration, the child started with the actual search tasks. Every search task within the sessions started on a very simple 'start page' (see Figure 4.7) with links to the websites (i.e. our study objects). Between each task, the child returned to this 'start page' with the browser's home button. The test instructor offered the tasks to the child verbally, to prevent the children from 'typing over' keywords instead of thinking about the formulation of the queries and the spelling of the words. For example, the test instructor asked the child: "Rembrandt was a famous Dutch painter and one of his most famous paintings is called 'De Nachtwacht'. Can you find the reason why he made this painting?" During task performance, the test instructor sat next to the child to reassure and help the child if necessary. After each task the test instructor held a short post-task interview. The test instructor did not steer children to

Figure 4.7: The 'start page' of the explorative study

use the search engine or to browse the main categories. The children could decide for themselves which search method they chose.

We made use of a within-subjects design, which means that all children conducted a search task on all websites. Four different search tasks were formulated for each website so that more web pages of the websites would be visited by the children than only the web pages that would be visited for one search task (see Appendix C)[2]. However, to be able to make comparisons between children that conducted the same search task on a website, each search task was conducted by a group of eight children. There were four different groups of children. The tasks were fact-based and not classroom related. In each group of children the websites were visited in a different order. Every child was offered a task about an animal, a task about arts, a task about sports and at the end a task about health (see Table 4.3). Every task belonged to the same domain and was formulated on the same level of abstraction. Also, the navigation path to a

2. The search task on Google was conducted first by all children and was for all children the same task: "Which day of the week will the Queen's birthday be on next year?"

Table 4.3: Task distribution over websites and over groups

	Group 1	Group 2	Group 3	Group 4
Task 1 Animals	Classic *(kangaroo)*	Classical play *(giraffe)*	Image map *(shark)*	(Wikikids) *(dolphin)*
Task 2 Arts or Music	Classical play *(Rembrandt)*	Classic *(Beethoven)*	(Wikikids) *(Mozart)*	Image map *(Van Gogh)*
Task 3 Sports	Image map *(hockey)*	(Wikikids) *(soccer)*	Classic *(basketball)*	Classical play *(gymnastics)*
Task 4 Health	(Wikikids) *(hay fever)*	Image map *(head lice)*	Classical play *(travel sickness)*	Classic *(braces)*

relevant content page was the same for the different tasks per website, to make them comparable to each other within the websites.

4.4.5 Collected data

The data collected consists of 16 hours of video and audio footage of the children's browser activities, eye movements and a video and audio display of the children in front of the screen during their research session. For each participant, we also collected data from the profile survey and the evaluations on the smiley-scales per search task. In total, we collected 96 pages of notes and comments made by the test instructor during the sessions.

4.4.6 Data analysis method

For the analysis of our data, we decided to use a top-down approach. First, we analyzed the quantitative data. After that, we analyzed the qualitative data to understand the processes and outcomes of the children's search performances and to explain the outcomes of the quantitative results.

Quantitative data analysis – pragmatic quality

The following variables per search task, that were measured during search performance, were analyzed:

1. Success in finding the relevant information
2. Help required from the test instructor
3. Number of events (clicks and submitted queries)
4. Deviation from the optimum navigation path
5. Amount of time needed to conduct the search tasks
6. Search strategy used (searching or browsing)

We determined the optimum navigation path by counting the number of clicks needed to find the right information on the websites in the most efficient way.

With these data, we could determine whether there were differences on the performance and evaluation scores between particular groups of children (for example, bet-

ween age groups), between the search tasks or between the visited websites. We could also check whether there were confounding variables, such as computer experience or prior subject knowledge.

Qualitative data analysis - pragmatic quality

For the qualitative data analysis, we studied the video and audio footage from the 32 sessions. We registered all problems and successes in navigation or comprehension during search performance. These observations were related to the system characteristics (such as layout, navigation tools or search engine characteristics) of the websites.

We did not impose categories for analysis before we started the qualitative data collection. Instead, we developed categories inductively after all the data had been collected by the test instructor, by categorizing and sorting all the positive and negative qualitative observations. We discovered that most developed categories could be assigned to one of the constituent processes in the Comprehension-based Linked Model of Deliberate Search (CoLiDeS) (Kitajima et al., 2000). This model divides the search process into several constituent processes: attending a web page, parsing a web page, focusing on an area, and selecting a relevant entry. Although this model only covers a browsing strategy, we experienced it to be useful for a keyword search strategy as well.

Analyzing affectivity

Identification of emotional expressions

Besides studying the pragmatic quality of the websites, we also studied the children's product experience. We choose to observe and analyze their expressive reactions towards the websites, because in our opinion this explorative method was best suited for the explorative nature of this study. We focused on verbal and non-verbal expressions of emotions, which we sorted into the emotion categories mentioned in Desmet's (2008) circumplex model of core affect. By emotional verbal expressions, we mean utterances that express some kind of emotional feeling towards the website, such as "Why can't I find anything on this website!" With emotional non-verbal expressions, we mean facial expressions or body posture that can also clearly express emotional feelings towards a website, such as a smiling face or an angry face. This type of verbal protocol with spontaneous utterances of children has been termed "voluntary think-aloud" by Donker and Reitsma (2004).

Reliability of the affective data collection

Five of the 32 sessions were analyzed by two independent evaluators (i.e. the author of this dissertation and a research assistant), from which the identified emotional expres-

sions were compared[3]. In total, 71% of the expressions that were identified by the evaluators as 'emotional' were evaluated as such by both evaluators. Differences about the remaining expressions were solved and based on that discussion agreement was reached about what expressions could be identified as 'emotions'.

After all sessions were studied and all expressions had been identified, the two independent evaluators assigned all expressions to the following categories: positive expressions (for example, "I like this duck!") versus negative expressions (for example, "What an annoying duck!") , the type of emotion that was observed and whether the emotional expression was caused by a pragmatic or a hedonic website characteristic. By measuring Cohen's kappa over the evaluations of the two independent evaluators, we could see whether the analysis of the data was reliable. For the labeling of positive and negative expressions, we found a kappa of .97 over all three types of website design. For the labeling of pragmatic versus hedonic expressions, a kappa of .91 was found over all three websites. This means that there was a high agreement between the evaluators about the labeling of the emotional expressions of valence and pragmatic versus hedonic reasons. For the labeling of the type of emotion observed, there was only fair agreement between the evaluators, with a kappa of .52 over all three types of website design. Therefore, no statistic analyses were conducted to explore whether there are differences in the type of emotions expressed by the children between the different types of website design.

Emotion types

The classification of the types of emotions that the children displayed was based on Desmet's (2008) wheel of product emotions (see Figure 4.2). Not all emotions mentioned in this wheel of product emotions are relevant for interaction with a website. Therefore, the emotions 'love', 'isolation', 'jealousy' and 'awaiting' were left out of consideration. Not all expressions could be classified by one of the product emotions mentioned in Desmet's wheel of product emotions. Three emotion types were added; relief (for example, "Finally, I found the answer."), frustration (for example, "How can I search for something on this website for heaven's sake?!") and tingling (for example, "Why is that duck swimming over there?"). After classifying all expressions, the following types of emotions were assigned to the expressions:

- Negative emotions: irritation, frustration, dissatisfaction, disappointment, contempt and boredom.
- Positive emotions: satisfaction, relief, joyfulness, fascination, tingling, astonishment, curiosity.

3. Manon van der Kooi assisted by the analysis of the children's emotional expressions in the context of her Master's thesis.

Hassenzahl (2004) says that he understands satisfaction as an emotional consequence of goal-directed product use. That is why we also classified satisfaction as an emotion type in this study.

Pragmatic versus hedonic emotional expressions

Emotional expressions were labeled as 'pragmatic' when the expressions were related to the effectiveness and efficiency of the search process, such as "I cannot find the answer" or "I finally found the page that I was searching for". Desmet and Hekkert (2007) also consider usability (pragmatic quality) to be a source of product experience, rather than a product experience itself.

Expressions were labeled as 'hedonic' when the expressions were related to aesthetics and experienced fun or pleasure. Examples of hedonic expressions are: "Wow, I saw the little robot move!" or "You can even play the piano as if it were a real piano!". Both examples express the joy that the children experienced with some playful elements on the websites, such as moving the robot and playing the piano.

4.5 Results – pragmatic quality

In this section, we will first discuss the quantitative data. After that, we will try to explain these outcomes by describing the most important observations within the qualitative data.

4.5.1 Control for confounding variables

Before addressing the search performance on the different websites in our study, we will first address whether there were differences between different groups of children. We conducted a one-way ANOVA to compare the means on several variables for different age groups, grade levels and gender.

There were no significant differences between the different age groups of children, concerning computer experience and Internet experience and prior knowledge of the subjects in the search tasks. There was one difference in the frequency of Internet use between grade levels ($F_{3,29} = 3.25$, $p = .036$)[4]; the higher the grade level, the more the children made use of the Internet. Concerning experience, there was one significant difference between genders ($F_{1,31} = 10.33$, $p = .003$); relatively more of the boys watched movies on the Internet.

4. The degrees of freedom for the numerator are the degrees of freedom for the between group and the degrees of freedom for the denominator are the degrees of freedom for the within group.

We did not find differences for the performance variables between different age groups, different grade levels or genders. We only found significant differences concerning the time children from different grade levels needed to conduct the tasks on two of the websites, the Classic website ($F_{3,28}$ = 5.74, p = .003) and the Classical play website ($F_{3,28}$ = 8.36, p = .000); the higher the grade level, the less time the children needed to conduct the tasks on these websites.

4.5.2 On which website did children best perform search tasks?

In this section, we will give an answer to the first main question: With which type of website do children perform the search tasks best? On each website, four different tasks were conducted within the four groups of children. These tasks were designed to be comparable in complexity and to require the same number of clicks within the optimum navigation path. Therefore, new performance variables were composed by combining the performance scores from the four tasks per website to one score for each of the websites.

A multivariate test between subjects did, however, show significant differences between the tasks for the Image map website concerning the number of clicks and for the Classical play website concerning the success scores. We decided to discard a task from data analysis when the scores of two or more of the pragmatic variables (time, clicks, success, etc.) deviated significantly from the scores of the other tasks, which was not the case with the task scores on the Image map website and the Classical play website.

Table 4.4: Percentages and mean scores of the pragmatic variables per website type

Pragmatic variables	Website type	Mean (N=32)	Std. deviation
Percentage of the children that found the right information	Classic	87%	.336
	Classical play	91%	.296
	Image map	41%	.499
Percentage of the children that required help	Classic	31%	.471
	Classical play	13%	.336
	Image map	63%	.492
Time (in sec.)	Classic	152.3 sec.	93.0
	Classical play	136.4 sec.	90.7
	Image map	256.7 sec.	73.5
Deviation from optimum navigation path (in amount of clicks)	Classic	3.2 clicks	3.2
	Classical play	1.7 clicks	2.8
	Image map	9.2 clicks	6.7
Difficulty evaluation scores (Likert-scale scores from 1 to 5; 1 = very easy, 5 = very difficult)	Classic	2.42	.87
	Classical play	2.15	.94
	Image map	3.88	.60

When we removed the results of these deviating tasks from the dataset, the direction of the results that are presented in this section - with these deviating tasks included - stayed the same. We will, however, take into account these differences between tasks when describing more qualitative results from our study.

We looked at the differences in performance between the websites by calculating a 'repeated-measures ANOVA' for each of the compound variables. For the analysis of the performance variables 'number of events', 'time' and 'deviation from the optimum navigation path' the data of all respondents were included. Also children who did not find the right information or required help from the test instructor in finding the right information were included. The reason for this is that we think that searching without finding the information and receiving help in finding the right information belong to a natural search process for children.

Table 4.4 shows that only 41% of the respondents could navigate to the right information on the Image map website, compared to 87% on the Classic website and 91% on the Classical play website (despite of the fact that one of the tasks on the Classical play website was significantly less successful than the other tasks on this website). There was a significant main effect of finding the right information on the websites, $F_{1.62, 50.20} = 16.14$, $p = .000$ The degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity[5]. Contrasts revealed that the respondents' success scores were significantly lower on the Image map website than the success scores on the other two websites, $F_{1, 31} = 20.67$, $p = .000$.

Also, significantly more respondents required help from the test instructor while working with the Image map website than while working with the other two websites, $F_{2,62} = 11.25$, $p = .000$ (see Table 4.4 for mean help-scores). For example, children required help with the correct spelling of a query, with navigating to the right category, with operating the tools or with scanning a content page.

The average amount of time needed to conduct the tasks was longest with the Image map website and shortest with the Classical play website, followed by the Classic website (see Table 4.4). There was a significant main effect of the total time needed to conduct the tasks on the websites, $F_{2, 58} = 19.99$, $p = .000$. Contrasts revealed that the time needed to conduct the tasks on the Image map website was significantly longer than the time needed to conduct the tasks on the other two websites, $F_{1, 29} = 27.61$, $p = .000$.

There was also a significant main effect of the average deviation from the optimum navigation path, $F_{1.46, 45.19} = 26.85$, $p = .000$. The degrees of freedom were corrected

5. The degrees of freedom used to assess the F-ratio are the degrees of freedom for the effect of the model and the degrees of freedom for the residuals of the model. The degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity, which is why the degrees of freedom are not round figures.

using the Greenhouse-Geisser estimates of sphericity[6]. Contrasts revealed that the average deviation from the optimum navigation path of the Image map was significantly larger than the average deviation using the Classical play website (see Table 4.4), $F_{1, 31} = 43.08$, $p = .000$. Also, the average deviation using the Classic website was significantly larger than the average deviation using the Classical play website, $F_{1, 31} = 4.69$, $p = .038$. Apparently, the design of the Classical play website supported efficient search behaviour. This might be caused by the images that were presented with the main category labels.

4.5.3 What type of website was evaluated as most difficult by the children?

As with the performance variables, we also composed variables for the evaluation of difficulty of the tasks for each website and calculated a repeated-measures ANOVA for these compound variables of difficulty evaluations. In general, the children evaluated the tasks on the Classical play website and the Classic website as more easy than the tasks on the Image map website (see Table 4.4). There was a significant main effect of the difficulty evaluation scores, $F_{2, 64} = 40.99$, $p = .000$. Contrasts revealed that the difficulty evaluation scores of the Image map were significantly higher than the difficulty evaluation scores of the Classical play website, $F_{1, 32} = 64.90$, $p = .000$.

4.5.4 Conclusion about pragmatic quality of the websites

The quantitative data results of pragmatic quality are very clear about the question on what type of website children perform the search tasks best. For each variable of pragmatic quality, the Classical play website had the highest scores and the Image map website had the lowest scores. The answer to research question chapter 4-1 is clear. Search performance was most effective on the Classical play website, because children were most successful in finding the right information on this Website and here they required the least amount of help from the test instructor. Search performance was also most efficient with the Classical play website, because the children needed the least amount of time and they deviated less from the optimum navigation path. The scores for the Image map website were exactly the opposite: search performance was least effective and efficient while using the Image map website. The children evaluated the Classical play website as easiest to work with and the Image map website as most difficult to work with. In conclusion, the Classical play website's pragmatic quality was highest, followed by the Classic website's pragmatic quality. The pragmatic quality of the Image map website was much lower than for the other two websites.

---

6. The degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity, which is why the degrees of freedom are not round figures.

4.5.5 Problems children experience with the websites

Why is pragmatic quality highest on the Classical play website and lowest on the Image map website? In other words; what problems did children experience while conducting the search tasks on the websites and to what characteristics of the websites did these problems relate? In this section, we will try to explain the quantitative results of the websites' pragmatic quality by addressing some important qualitative results of the study. Some of the problems that the children experienced with the Image map website will be described. These can explain the fact that children found the tasks on this website more difficult and preferred this website less than the other two websites.

Parsing problems

After a user attends a web page, the web page is parsed in several areas and the relevant area is focused on (Juvina & Van Oostendorp, 2008). We recognized different forms of 'parsing' in the children's search sessions. Some children did indeed look at all areas of a web page, before focusing on a relevant area (for example, on the main menu). Others only scanned a few items, before focusing on a particular area. Also, many children went straight to the relevant area with their eyes, without parsing other areas of the web page. They 'shortened' the parse process, because they had clear expectations about the page arrangement and looked at the area that was conventional for that item (for example, the conventional location for the search engine is at the top right corner of the web page). This was often the case on the Classic and Classical play website where items were placed at conventional locations. Most parsing problems were experienced using the Image map website. Children experienced problems with parsing, because items were placed at unexpected locations and the labels of main categories were only visible when 'scrubbing the screen', which will be described in the next section.

Mine-sweeping navigation metaphors

In their study on the usability of children's websites, Nielsen and Gilutz (2002) found that children were willing to indulge in mine-sweeping behaviour. However, we found exactly the opposite in this study. The home page of the Image map website presented a navigation metaphor in which the child navigated the page by 'sailing in a boat' and in which the child could visit different islands by clicking. However, many children did not understand this form of navigation. With the island 'Knowing', for example, there were two kinds of problems. The first problem was that this label 'Knowing' was too abstract. Children expected to find information here about almost everything. Further, the children had to 'scrub the screen' with the mouse to find the labels that belonged to these subcategory images, because they had no direct visible words explaining their

Figure 4.8: The sub page 'Knowing' on the Image map website without visible link labels

meaning as can be seen in Figure 4.8. In our study, this playful type of navigation proved not only to be ineffective for children, but also caused a lot of irritation.

Looping navigation style

The children often went back to pages they had already visited before, although they had not found the relevant information there. For example, when using the Image map website, children were confused about which main category was relevant for their search task. They often returned to the same main category 'Knowing' because they hoped they could find the information there although their first time visiting this category had not been successful. We also saw this 'looping' behaviour while children processed search results from the search engine. Bilal and Kirby (2002) reported the same results in their study on children's search behaviour. They found that most children had a "loopy" navigation style. They explain that this "loopy" style can be caused by children's lower cognitive recall, because the web imposes memory overload that reduces recall during navigation.

Home as 'comfort zone'

Most children went all the way back to the home page, when they started a new task within the same website. Navigating to a new page from a deeper page, was often too complicated for them. Starting from the beginning was more easy and clear. Problems arose when there was no clear home button or when the home button referred back to

an overview page for more target groups instead of the home page for children. Still, also when there were clear home buttons presented, children found it easier to use the browser's back button to go back to the home page, which is an inefficient way to go back. This search strategy to go back a couple of times or back to the home page was already mentioned by Chen (2003), who termed this strategy "going back to the comfort zone".

Search engines failing because of natural language queries

Many children used natural language when formulating their queries in a search engine, especially the younger children (8 years old). With the Classical play website and with Google, natural language querying did not cause any problems. However, the search engines on the other websites did not work well with natural language queries. The children often did not understand that the problems were caused by using a whole sentence, and tried to adjust the spelling of the words in their sentence. They did not think of bringing the query back to one keyword.

Spelling and typing

Spelling turned out to be a major obstacle for the children. Particularly with the interfaces that did not offer spelling correction. The children frequently asked the test instructor if their spelling was correct, as Druin *et al.* (2009) also found in their study. In our study most children gratefully made use of the spelling correction tool 'Did you mean' in Google. They seemed to be very experienced using this tool. Many children immediately clicked on the spelling suggestion after the search results were displayed. Typing also caused a lot of problems. Most of the children had to 'hunt & peck' for the right keys and did not notice when they made typing mistakes. Only two girls (ages 10 and 11) were able to use the touch-typing method. These results confirm the findings from Druin et al (2009) that familiarity with technology still has not allowed children to become proficient at typing. However, we do think that familiarity with Google allows children to overcome problems with spelling more and more.

Query suggestions

Two of the websites in our study (the Classic website and Google) offered query suggestions in a drop-down box while a word was being typed in the search box. Although children had their eyes on the keyboard while typing, children did quite often look at the screen while typing and many children took notice and made use of the query suggestions when offered. Some children used the query suggestions when they were not sure of the right spelling by checking whether the right keyword would come up. Others even used the query suggestions as 'type help' so that they had to type only a few letters.

For example, one boy only typed in the letters 'moz' on the textual interface and then clicked on the query suggestion 'mozart' that appeared in the drop-down box below.

Presentation of images with category links

Some children experienced problems with the search engine's results pages of the Classic website and the Image map website, because these children did not recognize the results as such. They thought that the summary or snippet was all the information there was to get and did not understand that they could click on one of the results to read more about the subject (see Figure 4.9). This problem did not happen with the Classical play website, because this website presented the search results with one sentence marked clickable and with a picture in front of each result. With this format, the children knew that they could click on the result for more information (see Figure 4.10). We saw the same positive effect of the use of images with categories or subcategories. When images were placed in front of subcategories (as was the case on the Classic website), the children recognized the categories as clickable more easily and scanned the list of categories by looking at the pictures.

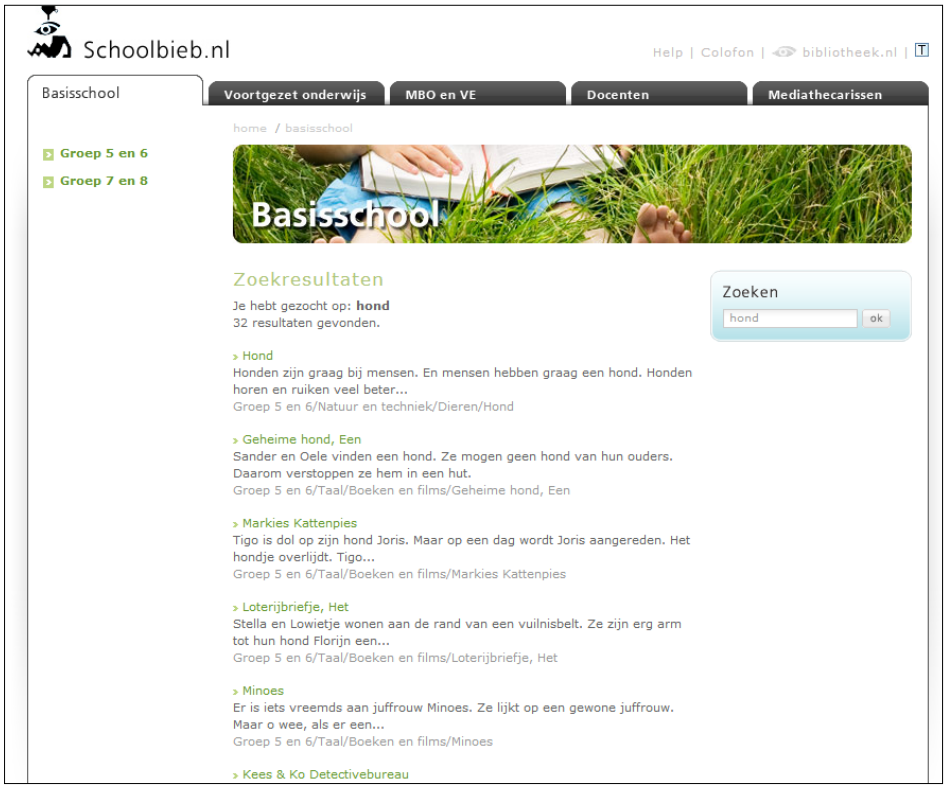


Figure 4.9: The search engine results of the Classic website without images in front of the search results

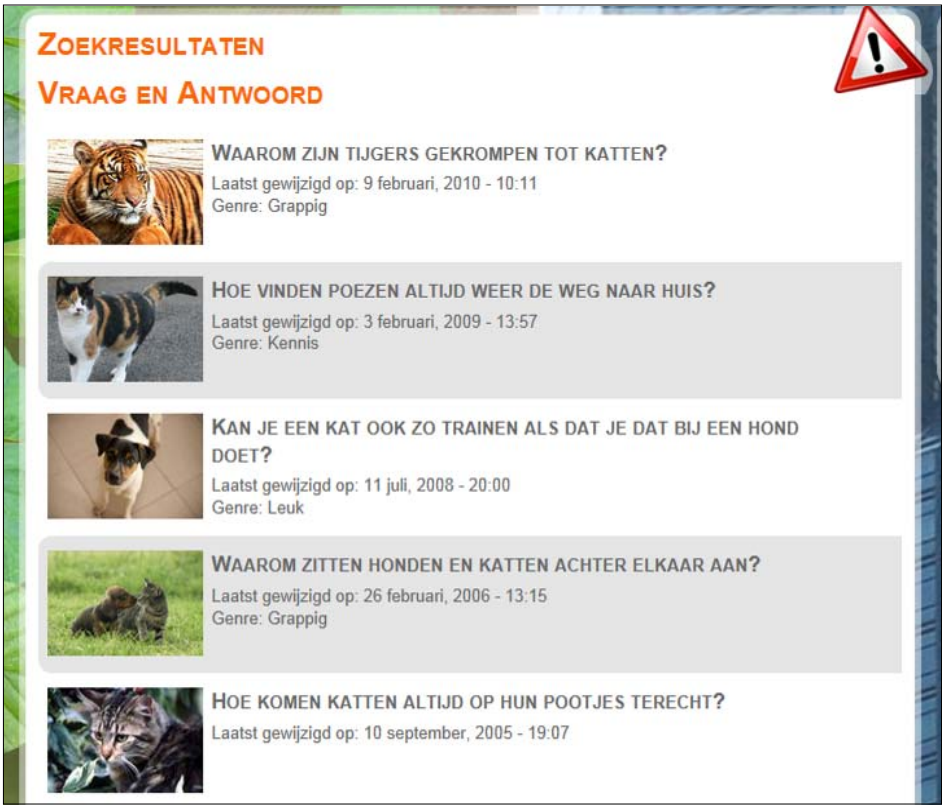


Figure 4.10: The search engine results of the Classical play website with images in front of the search results

Table 4.5: Problems experienced on the three different type of websites

Problems experienced with...	Classic website	Classical play website	Image map website
Parsing	-	-	X
Navigation (mine-sweeping)	-	-	X
Looping navigation	-	-	X
Going 'back to home'	X	-	X
Natural language querying	X	-	X
Presentation of search engine results	X	-	X

When looking at the type of problems that were addressed in Table 4.5, we see that all types of the problems discussed emerged while children were working with the Image map website and that none of the problems discussed emerged at the Classical play website.

4.6 Results – affective user experience

With the ranking of the websites for future use at the end of each session, the children almost unanimously ranked the Image map website as least preferred for future use. After Google, the Classic website was chosen as the most preferred website for future use, closely followed by the Classical play website (see Table 4.6). Probably, the main cause for these preferences is the fact that the actual and perceived pragmatic quality of the Image map website was much lower than the pragmatic quality of the Classic and Classical play website. However, how would children have ranked the three websites when only considering non-instrumental qualities of the websites? Is the perceived hedonic quality of the websites different from the pragmatic quality of the websites?

In this study, the affective side of user experience was measured by analyzing children's verbal and non-verbal emotional expressions during search performance for all tasks conducted on the Classic, Classical play and Image map website. Comments expressed during the post-task interviews were left out of consideration, because the affective experiences that were uttered, were not uttered spontaneously, but in retrospect. These involuntary comments did not add value to the emotions expressed during task performance (voluntary think-aloud). Often, these utterances were a repetition of their spontaneously uttered emotions during task performance or the utterances tended to be socially desirable in contrast to the emotions expressed spontaneously. In total, 273 emotional expressions were identified; 214 verbal expressions, 34 non-verbal expressions and 25 expressions were both verbal and non-verbal at the same time (e.g. a child looks angry while he complains that he cannot find the right page). Because of the low number of verbal expressions and combination expressions, all expressions will be treated as the same type of 'emotional expressions'. In some cases there were no emotional expressions identified at all. This was the case for 13 children on the Classic website and for 12 children on the Classical play website. These children conducted the

Table 4.6: Final ranking of the websites for future information search tasks

Website type	**Rank1**	**Rank2**	**Rank3**	**Rank4**	**Rank5**
(Google)	**28**	2	0	3	0
Classic website	2	**13**	**9**	3	0
Classical play website	2	**9**	5	5	5
(Wikikids)	1	3	**8**	**11**	1
Image map website	0	2	2	2	**26**
Subtotal	33	29	24	24	32
Missing*	0	4	9	9	1
Total	33	33	33	33	33

* Some children were not able to give websites a rank 2, 3 or 4. They were able to evaluate which website was best and which website was worst in their opinion, but they did not rank the websites 'in between'. This explains the 'missing values' with the ranks in the table.

search tasks very quickly, without expressing any feelings at all. It is remarkable that all children uttered one or more emotional expressions while working on the Image map website.

When comparing the total of emotional expressions, a major difference can be seen between the total number of emotions expressed between the Image map website and the other two websites. In total, 52 emotional expressions were identified with the Classic website, 40 emotional expressions were identified with the Classical play website and 181 emotional expressions were identified with the Image map website.

4.6.1 Valence of the emotional expressions

The first question about the emotional expressions identified is whether there was a difference in valence between the three websites. According to the Pearson's chi-square test, there was a significant association between the type of website and the number of positive and negative emotional expressions identified χ^2 (2) = 7.89, $p < .05$. This association seems to be mainly caused by the distribution of number of positive and negative feelings experienced with the Classic website. It seems that the children experienced relatively more negative feelings than positive feelings on the Classic website (see Table 4.7).

4.6.2 Type of emotions

The next question concerning the emotions experienced is what types of emotions children displayed while using the websites. When looking at the types of emotions expressed that are presented in Table 4.8 and Table 4.9, it can be seen that most displayed emotions were satisfaction and dissatisfaction. Unfortunately, because there are too many cells with an expected count lower than five and because there was only

Table 4.7: Number of positive and negative emotional expressions for the three websites

Valence of emotional expressions (N=273)		Positive	Negative	Total
Classic website	Count	15	37	52
	% within website	28.8%	71.2%	100%
	Std. Residual	-1.8	1.7	
Classical play website	Count	19	21	40
	% within website	47.5%	52.5%	100%
	Std. Residual	.1	-.1	
Image map website	Count	92	89	181
	% within website	50.8%	49.2%	100%
	Std. Residual	.9	-.9	
Total	Count	126	147	273
	% within website	46.2%	53.8%	100%

fair agreement between the two evaluators concerning the types of emotions that could be assigned to the utterances, no further statistical analyses could be performed with the types of emotions expressed.

The children expressed more negative than positive emotions while working with the Classic website. The types of negative emotion expressed most while working on this website were dissatisfaction and disappointment. An example of an expression of dissatisfaction is "I can't find it" and an example of an expression of disappointment is "Hmmm… no, it doesn't work without the search engine." Most experienced positive emotion was satisfaction on this website. Emotions such as astonishment, curiosity, tingling and fascination were not experienced on this website. This finding suggests that the design of the Classic website did not have a positive effect on the children's perception of non-instrumental qualities of the website.

On the Classical play website satisfaction and dissatisfaction were also the two most displayed positive and negative emotions. However, with the Classical play website, the children displayed more satisfaction than dissatisfaction, while this was the other way around with the Classic website. Also different from the Classic website was that a few children also experienced a tingling emotion (for example, "Oehhh, games!") and one child was even astonished by something on the Classical play website (for example, "Hey, it's over here!"), which was not experienced by the children on the Classic website. These results suggest that the playful design of the Classical play website had a positive effect on children's perception of non-instrumental qualities of the website.

Table 4.8: Frequencies of positive types of emotions on the different websites

Positive emotions	Classic	Classical play	Image map	Total
Satisfaction	9 (17%)	11 (28%)	25 (14%)	45 (16%)
Joyfulness / fun	1 (2%)	2 (5%)	26 (14%)	29 (11%)
Tingling	0	4 (10%)	13 (7%)	17 (6%)
Relief	5 (10%)	1 (3%)	10 (6%)	16 (6%)
Fascination	0	0	10 (6%)	10 (4%)
Astonishment	0	1 (3%)	6 (3%)	7 (3%)
Curiosity	0	0	2 (1%)	2 (1%)

Table 4.9: Frequencies of negative types of emotions on the different websites

Negative emotions	Classic website	Classical play	Image map	Total
Dissatisfaction	19 (37%)	9 (23%)	36 (20%)	64 (23%)
Disappointment	10 (19%)	5 (13%)	16 (9%)	31 (11%)
Irritation	6 (12%)	4 (10%)	16 (9%)	26 (10%)
Frustration	1 (2%)	3 (8%)	14 (8%)	18 (7%)
Contempt	1 (2%)	0	6 (3%)	7 (3%)
Boredom	0	0	1 (1%)	1 (1%)

The children experienced more fun on the Image map website than on the other websites: for example, "How funny", when feeding the fish in the fishbowl and "This is fun!", while playing the piano. Still, the most negative type of emotion displayed on the Image map website was dissatisfaction. This might be explained by the fact that children could not navigate effectively and efficiently on the Image map website. Also, disappointment was a frequently displayed emotion on this website, because often children expected to find information on the island 'Knowing', but were disappointed by the fact that they could not find the information on this web page.

4.6.3 Pragmatic versus hedonic emotional expressions

The fact that the children displayed more negative feelings with the Classic website than with the other two websites and that children displayed most feelings of satisfaction with the Classical play website, still does not tell us much about the causes of these feelings. Therefore, the two evaluators also judged whether an emotional expression was uttered because of a pragmatic cause (concerning the effectiveness and efficiency of the search process) or because of a hedonic cause (concerning the beauty of the design or the fun experienced by a playful element on the website).

The most remarkable result was that the greater part (70%) of the emotional expressions were caused by pragmatic characteristics (see Table 4.10). There was a significant association between the type of website and whether the emotional expression was labeled as pragmatic or hedonic χ^2 (2) = 36.51, $p < .01$. For both the Classic website and the Image map website, the standardized residual was significant for the hedonic emotional expressions. And for the Classic website this was also significant for the pragmatic emotional expressions. When using the Classic website significantly more emotional expressions than expected were caused by pragmatic reasons and significantly fewer emotional expressions than expected were caused by hedonic reasons. When using the Image map website, significantly more emotional expressions than expected were caused by hedonic reasons.

With all three websites, most pragmatic emotional expressions were negative. The children mostly expressed their feelings when they experienced problems within the information searching process. With the hedonic emotional expressions it was the other way around. Concerning hedonic elements, such as aesthetics and playful elements, children expressed more positive feelings than negative feelings.

Relatively, more emotional expressions with the Image map website were based on hedonic causes than with the other two websites. It seems that the aesthetic design of the Image map website and the fun experience by the playful elements on this website stimulated children in expressing emotions. No emotional expression about the Classic website was based on a hedonic issue at all.

Table 4.10: Frequencies of pragmatic and hedonic based emotional expressions on the different websites

Hedonic versus pragmatic reason for emotional expressions (N = 273)		Hedonic reason	Pragmatic reason	Total
Classic	Count	0	52	52
	Expected Count	15.6	36.4	52.0
	% within website	0%	100%	100%
	Std. Residual	-4.0	2.6	
Classical play	Count	7	33	40
	Expected Count	12.0	28.0	40.0
	% within website	17.5%	82.5%	100%
	Std. Residual	-1.4	.9	
Image map	Count	75	106	181
	Expected Count	54.4	126.6	181.0
	% within website	41.4%	58.6%	100%
	Std. Residual	2.8	-1.8	
Total	Count	82	191	273
	Expected Count	82.0	191.0	273.0
	% of Total	30%	70%	100%

4.6.4 Conclusion about affective user experience

Working with the Image map website evoked most emotional utterances expressed by the children in our study. Two-thirds (181 of the 273 emotional expressions) of the uttered emotions were about the Image map website.

The Classic website evoked mostly negative emotions, because of low instrumental qualities. Children experienced usability problems on the Classic website, such as finding the right main category, finding the way back to the home page, or finding the right answer while reading a content page. Also the positive emotions expressed while using the Classic website were based on instrumental qualities, such as satisfaction or relief in finding the answer to the search task.

Most emotional expressions concerning the Classical play website were also based on pragmatic reasons, mostly uttered in satisfaction or dissatisfaction/disappointment in finding or not finding information on the website. However, there were some emotional expressions based on non-instrumental qualities of the Classical play website, such as fun, tingling or astonishment about the playful elements on the website, such as the games and movies.

The Image map website evoked more emotional expressions than the other two websites. The number of positive and negative emotional expressions was almost equal. Satisfaction and dissatisfaction were the most frequent type of emotions uttered. Although most emotional expressions concerning the Image map were based on pragmatic reasons (106 expression), the number of emotional expressions based on hedonic

reasons was also large (75 expressions). Apparently, the playful design of the Image map website evoked a lot of hedonic, emotional expressions, such as joyfulness and tingling about the "fun stuff", but also irritation about "all the distracting animations".

4.6.5 Relation between perceived pragmatic and hedonic quality

The third research question is about the relation between perceived pragmatic and hedonic quality and the role of design characteristics. When we compare our quantitative results for perceived pragmatic quality with the identified emotional expressions, we do not see the same pattern. The Classical play website, that had the highest scores on pragmatic quality, evoked a lot fewer emotional expressions than we expected; only 15% of all expressed emotions. The expressed emotions concerning the Classical play website were also not predominantly positive. In other words, children did not express a lot of positive emotions while their search performance was efficient and successful.

The children did, however, express emotions when their performance was inefficient and unsuccessful as can be seen by the emotions expressed concerning the Image map website. The Image map website had the lowest scores on pragmatic quality and evoked many more emotional expressions, both positive and negative. The negative emotions were predominantly based on pragmatic quality. It seems that there is a stronger relation between low pragmatic quality and negative emotional expressions than between high pragmatic quality and positive emotional expressions.

The playful characteristics of the Image map website, however, evoked more emotional expressions concerning perceived hedonic quality than the other two websites in this study. There seemed to be a stronger relation between high perceived hedonic quality and the number of positive emotions expressed than between low perceived hedonic quality and the number of negative emotions expressed.

The final ranking of the three websites shows another picture than the analysis of perceived pragmatic and hedonic quality showed, as can be seen in Table 4.11. All children ranked Google as the most preferred website for future information search tasks. For this chapter, Google was left out of consideration as described in Section 4.1. This might have blurred the ranking of the Classic and the Classical play websites by the children. They were very confident about the ranking of the 'best' and the 'worst' website, but they were uncertain about the ranking of the websites 'in between', as will be discussed further in Section 4.7.1.

Table 4.11: Ranking of the websites concerning the CUE-components according to our study

Objective pragmatic quality *(amount of required time, clicks and success)*	**Perceived hedonic quality** *(number of positive expressed emotions)*	**Overall evaluation** *(ranking of the websites for future information search tasks)*
1. Classical play	1. Image map	1. Classic
2. Classic	2. Classical play	2. Classical play
3. Image map	3. Classic	3. Image map

4.7 Discussion

In the study reported in this chapter, we explored the influence of playfulness added to visual and navigation design on children's search behaviour and their attitude towards informational websites designed especially for children. The effects of different design characteristics on both usability and subjective judgments were analyzed. Lindgaard and Dudek (2003) report in their study about usability of interactive systems with adults that subjective judgments of the use of systems were mostly based on efficiency and effectiveness of system usage. While trying to measure subjective judgments based on non-instrumental qualities of a website, we also found that most emotions uttered by children were based on efficiency and effectiveness of the websites.

Many problems that children experienced concerning usability of the websites have also been reported in previous studies, such as problems with formulating search queries, correct spelling, natural language querying (Druin et al., 2009), problems with finding the right category because of abstract top headings, or problems with mine-sweeping main and sub category pictures. These are all issues based on instrumental qualities of the websites. In this study, we also tried to shed more light on the influence of non-instrumental qualities on children's evaluation of a website and to the question whether emotional reactions to the interaction with a website are mostly influenced by instrumental or non-instrumental qualities.

This study revealed some interesting results compared to current literature. Apparently, for children, perceived pragmatic and hedonic quality did not have such a large influence on the overall evaluation of an interactive product as reported for adults by Thüring and Mahlke (2007) in their CUE-model. We measured the overall evaluation by asking which of the websites children preferred most for future information search tasks (see Table 4.6). We would have expected children to rank the Classical play website as most preferred for future use, because working with this website was most effective and efficient. However, the Classic website was preferred slightly more than the Classical play website. Some children might have ranked the Classic website higher than the Classical play website, because they thought that this is expected of them by their parents or teacher.

Even more surprisingly is the opposite pattern between perceived pragmatic and hedonic quality, especially concerning the Image map website. Apparently, it seems that perceived pragmatic quality was more important for children than hedonic quality for the overall evaluation of the websites. We found that relatively more emotional expressions concerning the Image map website were based on hedonic reasons than concerning the other two websites. It seems that the aesthetic design of the Image map website and the fun experience by the playful elements on this website stimulated children in expressing emotions. However, this had no positive influence on overall evaluation. In contrast, no emotional expression about the Classic website was based on a hedonic issue at all, but overall evaluation was highest for the Classic website.

It seems that hedonic issues did provoke emotional expressions, but these were not of importance for children when evaluating the system for future information search tasks. In other words, classic aesthetics did not provoke positive hedonic reactions from children, but expressive aesthetics did (Lavie & Tractinsky, 2004). In addition, the influence of aesthetics on overall appraisal of a system seems very small with children. Pragmatic quality of a website had much more influence on children's overall appraisal of the interfaces. Thüring and Mahlke (2007) also report that the effect of usability in their study was greater than visual aesthetics for both valence and arousal. This is not in line with the results of Read and MacFarlane (2006) that proved that the major factor in a child's decision about whether they want to use an interactive product again is how much fun it was instead of ease of use. Based on our study, we expect that the major factor in a child's decision about future use of an interactive product is ease of use instead of how much fun it was. The experiment that we will report in Chapter 5 should provide evidence for this hypothesis.

4.7.1 Children's performance using Google and Wikikids.nl

We found that the children's prior experience with Google and the influence of this experience on their general search behaviour should not be underestimated. Compared to the other websites, the children's search performance was most effective and efficient using Google. They judged the Google tasks as easiest to perform. They also evaluated Google as most preferred to use for future information search tasks. Our findings concerning the children's performance on Google were also more positive than Druin, Foss, Hatley, Golub, Guha, Fails and Hutchinson (2009), because in our study children did take advantage of the Google-tools, such as spelling corrections and query suggestions. The most important problem with Google for the children was judging the relevance, suitability and reliability of search results. This is one of the most important reasons why search interfaces are designed especially for children, with access to a content directory of information especially selected for children.

The children's search performance on Wikikids.nl was left out of consideration for the results reported in this chapter. In short, the children experienced a lot of problems using this website. This was not clearly because of certain considered design characteristics (as was the case with the playful navigation design of the Image map website), but because of overall non-user friendly web design. For example, the home page included too much text for children and it was unclear for the children where to look for the main categories on the home page. Therefore, this website type was not represented in the experiment that is reported in Chapter 5.

4.7.2 Evaluation of used methods

Another goal of this study was to explore methods that can be used in research on children's digital search behaviour. Logging the activities of the children and measuring pragmatic variables proved to be an effective method in this study. The method was unobtrusive for the children and could run alongside the other evaluation methods, as reported by Markopoulos, *et al.* (2008). Using the eye-tracker was a valuable method, because children's screen activities and spoken comments could be very precisely combined with the area on the screen on which the children focused their attention. For research on the effects of overall interface design, children's screen activities in combination with their spoken comments without their eye movements, might provide enough insight into children's interaction with the interface. Therefore, we doubt whether it is worth the effort and expense of using an eye tracking system with this type of research. Where the research goal is to identify user problems in relation to specific design characteristics, Elling, Lentz and De Jong (2012) argue that eye tracking data is even of more value than users' verbalizations. They evaluated different methods for testing web usability with adult users and they found that the combination of concurrent think-aloud with observational methods such as eye tracking, led to the most complete overview of user problems, thanks to the added value of the interaction between both sources of information.

We decided to conduct post-task interviews after each search task instead of using a think-aloud protocol, to prevent extra cognitive load on the children during the search process itself. However, the results of the post-task interviews did not add value to the observations made by logging the screen activities and making video and audio recordings of the children during search performance. Children did not report many problems or successes and had trouble reflecting on their own search process, as was also reported by Baauw and Markopoulos (2004).

We used smiley scales to measure children's appraisal scores of pragmatic quality of the interfaces. We did not find a clear tendency of children to give extremely positive ratings as Read *et al.* (2002) did. This might be caused by the fact that children were asked about how difficult the task was, so not only the interface was evaluated, but

also the interaction with this interface. Neither did the children report problems with the neutral position on the scale which had a face with as straight-line mouth, as Read (2008) found. Moreover, there was clearly a tendency to choose this neutral position on the smiley-scale, which had the label 'normal'. Apparently, children were not very agitated by the tasks and interfaces and often evaluated them as "normal; not very easy, but also not very difficult".

A more objective method to measure children's emotional experience than asking about their emotional feelings, is to observe their expressive reactions towards the websites. By observing children's spontaneous utterances (termed "voluntary think-aloud" by Donker & Reitsma, 2004), we hoped to get a clear picture of children's product experience. This method proved to be very helpful in our qualitative comparison of children's product experience between the three different websites. Children did spontaneously utter emotional reactions towards the websites. It turned out that most expressive reactions were attributed to usability (pragmatic quality) of the websites and not to non-instrumental qualities, such as aesthetics and playful design. However, as product experience is a multi-faceted phenomenon that involves manifestations such as subjective feelings, behavioural reactions, expressive reactions, and physiological reactions (Desmet & Hekkert, 2007), future research should cover more of the facets of product experience to validate the picture presented here.

Initially, the children ranked five websites instead of the three websites, as described in Section 4.1. They had no trouble at all ranking the websites that they found best and they found worst for future use. Almost all of the children ranked Google as the best website and the Image map website as the worst website. They had much more trouble ranking the websites 'in between' the best and the worst. Read *et al.* (2002) also reported this problem that children could easily rank the items that they found the best or the worst, but that it was much more difficult to assign items to the ranks 'in between'. These ranks were often not assigned at all.

4.7.3 Controllability of providing help

Contrary to the strict think-aloud of Ericsson and Simon (1993), it is the view of Markopoulos *et al.* (2008) that getting children to verbalize their thoughts depends on how much the test instructor can make them feel confident and relaxed. In their opinion it is better to sit next to the child to make him feel at ease and be there to help when appropriate. After conducting this study, we could not agree more with Markopoulos *et al.* (2008). We also experienced that it is very important to make children feel at ease and win their trust. Children often asked for confirmation whether their actions were correct. Some children required help from the test instructor directly from the beginning of the search process. They just needed a small push to get the process started and to win confidence. Otherwise, they would not interact with the system at all.

We do realize that it is very difficult to control the amount and the manner in which help is offered to the children. However, by analyzing the amount of help required by the children, we could also analyze the problems children experienced with the different websites. Also, in future research we cannot and will not avoid help and social interaction with the children, because of the reasons mentioned before. In experiments, this will make it a challenge to reach experimental rigidity and validity. By analyzing the amount, but even more important, the type of help offered (for example, help with spelling, navigating, operating navigational tools, etc.), we might make this help-factor more controllable.

4.7.4 Future research

This study demonstrates some of the effects of playful interface design on children's performance and affective responses. However, more empirical research is required to find evidence for the effects suggested by this explorative study. For example, it seems that adding images to textual category labels has a positive influence on search performance, as suggested in Section 4.5.2.

Based on this explorative study, it is not possible to deduce certainties about the difference between children's search performance by browsing categories and by searching with a search engine. On both the Classic and the Classical play website almost all children used the search engine (29 of 32 children), which makes a comparison between searching and browsing invalid. And almost all children (29 of 32 children) browsed the categories on the Image map website (see Figure 6.7), because the search engine was hard to find on the Image map website in the explorative study. Therefore, it was impossible to make a valid comparison between searching and browsing based on the results in this study.

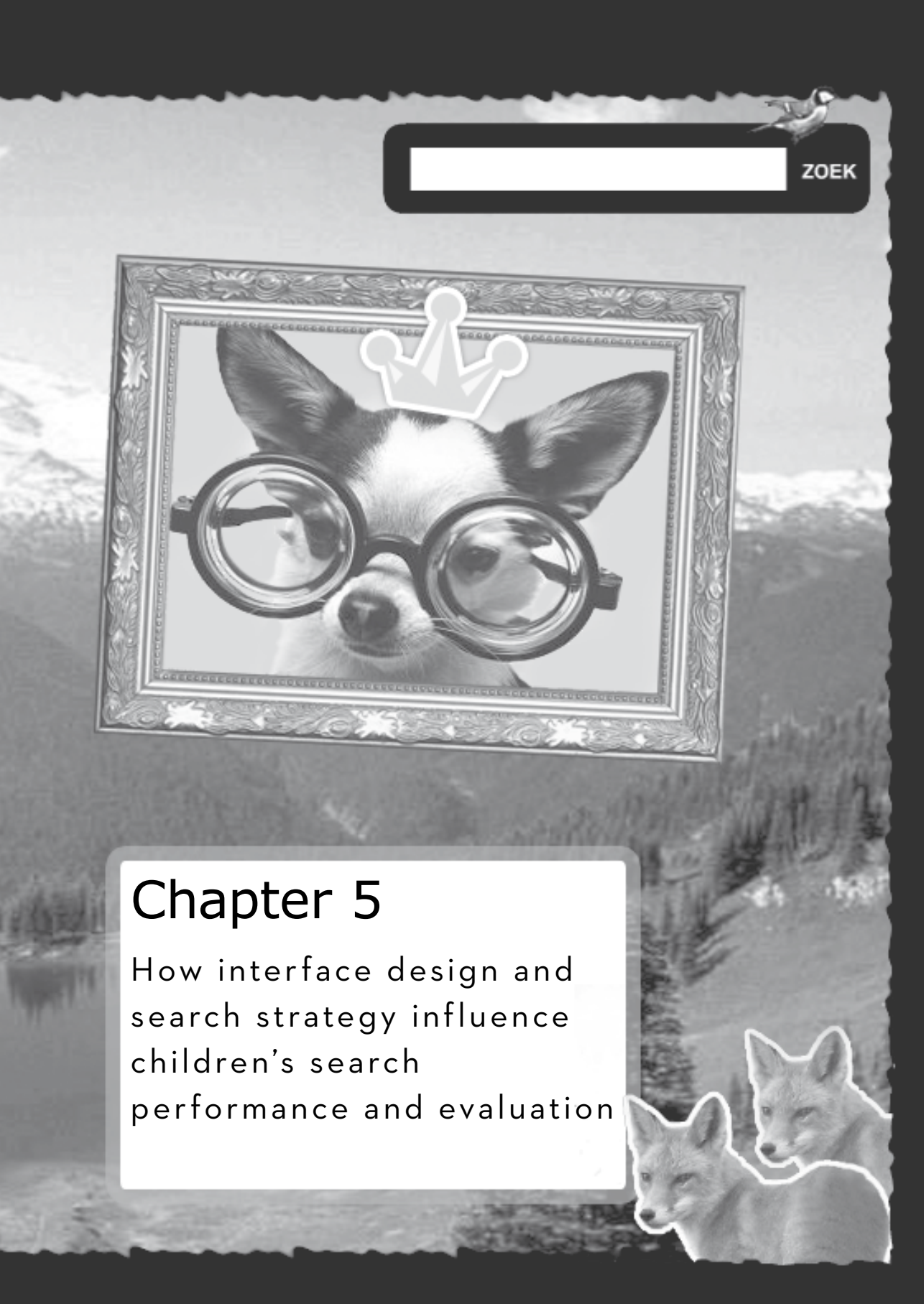
However, because of the fact that children's search performance was most effective and efficient using Google and because of the fact that most children used the search engine on the other websites (if they were able to find the search engine), we believe that searching is more effective and efficient for children than browsing categories. Future research should further test this hypothesis. We will test this hypothesis in the experiment reported in Chapter 5.

In the 1980's, Malone published guidelines for designing for fun (Malone, 1984). Also, more recently, there has been an interest in this aspect of software for children (Read et al., 2002). Although our study shows that playful design did provoke positive emotional expressions, it seems that playful design did not have a large influence on the children's overall appraisal of the informational interfaces. Usability of informational websites seemed to be much more important for the children than non-pragmatic qualities, such as fun and beauty. More research is needed to further investigate the effects of playful

visual design and playful navigation design on children's search performance, emotional responses and final appraisal of the interfaces.

We only used one method to measure the children's emotional responses to the different interface designs and one method to measure their final appraisal of the interfaces. Future research should apply more methods to measure effects of non-instrumental qualities on children's emotional responses and appraisal. In our experience, voluntary think aloud works better than post-task interviews. For measuring children's perceptions of pragmatic and hedonic quality and their appraisal of constructs such as beauty and fun, existing methods with adults (Van Schaik & Ling, 2008) should be tested with or adapted to children.

While the design of more playful informational interfaces for children is in development, more knowledge is needed about the real effects of these design efforts on children's search performance and attitudes towards these interfaces. Therefore, we conducted an experiment, to empirically test the suggested effects of playful visual and navigation design from this explorative study. The results of this experiment will be reported in Chapter 5.



Chapter 5

How interface design and search strategy influence children's search performance and evaluation

5.1 Introduction

There is a trend in digital media for children to design digital products that are 'cool' and 'playful'. Part of taking a 'playful' approach in designing digital products for children is to create age-appropriate graphics, or graphics that children can relate to (Meloncon, Haynes, Varelmann & Groh, 2010). In the corpus study of 100 informational websites for children that we reported in Chapter 3, we recognized this playful design approach in many of the interfaces we analyzed. More specifically, we identified three types of interface designed for children, 1) the Classic design type, 2) the Classical play design type, and 3) the Image map design type. In the explorative study reported in Chapter 4, we explored the possible effects of these different design approaches of informational websites on children's interaction with these interfaces and on children's affective responses towards these interfaces. This explorative study was not designed as an experiment and empirical research is required to test whether the effects suggested by the results of the explorative study can be validated. Therefore, in this chapter, we will report the results of an experiment on the real effects of these different design approaches of children's informational interfaces on children's search performance and on children's attitudes towards these interfaces. The second important objective of this experiment was to explore the effects of children's use of a search engine on the children's search performance. Conducting an experiment by letting children interact with digital interfaces was a big challenge. However, measuring children's affective responses towards these interfaces was an even greater challenge, as will be described in this chapter.

5.2 Theoretical background

5.2.1 Children's informational interface design

Interactive products for children can be classified as entertainment, educational and enabling products (Markopoulos, Read, MacFarlane & Hoysniemi, 2008). Websites for children as a specific group of interactive products can also be classified in these three genres. Most websites for children are aimed at entertaining children, for example by providing computer games. For our study with children's informational websites, both educational and enabling websites are relevant, because most informational websites are educational and search engines that help children in finding relevant information can be classified as enabling.

Researchers propose some guidelines for children's web design (Nielsen & Gilutz, 2002; Meloncon, et al., 2010). Most of these guidelines were tested and validated with children, but many of the guidelines are not specifically aimed at children and are similar to standard Web design practices for adult websites. In a large corpus study with

children's informational websites, which we reported in Chapter 3, we identified current design conventions for children. This study also showed that designers of children's websites often follow general Web design guidelines. A closer look at the data in this study did reveal three categories of informational websites especially designed for children. The first category is a 'Classic design type' in which the layout of the pages is kept minimal and the design is aimed at simplicity, consistency and focus. We called the second category the 'Classical play design type' in which a classic design approach for the navigation structure is combined with a playful, visual design approach. More effort is spent on the design of graphics, colors and games (Meloncon et al., 2010). The third category was called the 'Image map design type' in which no classic Web design characteristics are used. The visual design and navigation structure on the websites of this type are based on Image maps that incorporate objects or locations that children know from real life or from fiction. Children can explore this tableau of real life or fictional objects, which makes information search a playful experience (Meloncon et al., 2010). This Image map web design can be compared to 'spatial metaphors', which can be employed to represent information visually, using the universe, the solar system, galaxies, and so on through which the user navigates to locate information (Chen, 2006).

In their study to develop a visual taxonomy for children, Large, Beheshti, Tabatabaei, and Nesset (2009) emphasized the importance of movement and color in any visualization designed for children. They argue that "such characteristics do not necessarily influence positively the effectiveness of a taxonomy, but the affective reaction of users, and especially of children, that should never be underestimated. If the presentation is not interesting and fails to catch the attention of users, it is unlikely to invite their repeat visits. It also might be argued that intrinsic to visualization schemes is the ability to provoke interest and even fun" (p. 1818).

5.2.2 Two search strategies: keyword searching and browsing

At the beginning of the Internet era a general assumption was made by researchers that browsing-oriented search tools, relying on recognition knowledge, were better suited to the abilities and skills of children than keyword search tools. The argument was that browsing imposes less cognitive load on children than searching because more knowledge is needed to retrieve terms from memory when searching than simply to recognize terms offered when browsing (Bilal, 2000, 2001, 2002; Borgman, Hirsh, Walter & Gallagher, 1995; Large & Beheshti, 2000; Large, Beheshti, & Moukdad, 1999; Schacter, Chung & Dorr, 1998; Bilal & Watson, 1998).

Schacter *et al.* (1998) found that with both highly specific and vague search tasks, children sought information by using browsing strategies. In their research on children's internet searching on complex problems with thirty-two children between the ages of 10

and 12 years, they reported: "Children are reactive searchers who do not systematically plan or employ elaborated analytic search strategies" (p. 847).

Bilal (2000) found in her research on the use of the Yahooligans! Web Search Engine that most of the children (she observed twenty-two children between the ages of 12 and 13 years) used keyword search. Only 36% of the searches were performed by browsing under subject categories. This finding may have been affected by the type of search task that was given in this research: a fact-driven query that automatically stimulated children to use keyword search instead of browsing the categories.

Revelle, Druin, Platner, Bederson, Hourcade and Sherman (2002) report on the development of a visual search interface to support children in their efforts to find animals in a hierarchical information structure. To examine searching and browsing behaviour, 106 children (aged 5 through 10) participated in an experiment on this visual search interface. The researchers found that: "(…) even young children are capable of efficient and accurate searching. With the support of a visual query interface that includes scaffolding for Boolean concepts, children can use a hierarchical structure to perform searches and construct search queries that surpass their previously demonstrated abilities with the use of traditional search techniques" (p. 56).

By tracking the web logs of The International Children's Digital Library (ICDL), Druin (2003) found that, of 60,000 unique users between the ICDL's launch in November 2002 and September 2003, approximately 75% of the searches used category search (browsing), 15% used place search (by selecting a place using a world interface) and just over 10% of the searches used keyword search.

Hutchinson, Bederson and Druin (2006) found that children are capable of using both keyword search and category browsing, but that they generally prefer and are more successful with category browsing. Hutchinson *et al.* (2006) explain this finding in relation to children's 'natural tendency to explore'. Young children tend not to plan out their searches, but simply react to the results they receive from the Information Retrieval system. Generally, their search strategies are not analytical and do not aim precisely at one goal. Instead, they make associations while browsing. This is a trial-and-error strategy.

In our explorative study (Chapter 4), we did not control the use of the search strategies and therefore, we could not make a valid comparison between children's search performance using a search engine and their search performance by browsing the categories. However, we did see that children had a lot of experience in using Google to search for information and that they were more successful using Google in comparison to their search performance on the websites that were especially designed for children. Also, most children preferred to use a search engine over browsing main categories. Some children were even irritated when they could not find a search engine, which was often the case on the Image map website in the explorative study.

It is clear that research results are very diverse when it comes to search strategies used by children. The results seem to depend on the type of interface used in the studies and the type of search task that is given to children. However, the trend in the literature is that browsing is more suited for children than using a search engine.

5.2.3 Difficulties with keyword searching and browsing

Formulating a search query might be difficult for children because they have little knowledge to base 'recall' on (Borgman et al., 1995; Hutchinson, Druin, Bederson, Reuter, Rose & Weeks, 2005). Besides, for searching relevant documents using keyword search, correct spelling, spacing and punctuation are needed. Children have difficulty with spelling and often make spelling errors (Borgman et al., 1995; Druin, Foss, Hatley, Golub, Leigh Guha, Fails, 2009). Deciding on a single keyword is also difficult for a child because children tend to use a full natural language query, especially with complex search tasks (Marchionini, 1989; Druin, et al., 2009). In our explorative study (see Section 4.5.5), incorrect spelling and natural language querying did also cause problems for children when they used the search engines. Most of the search engines did not deliver any results when queries were spelled incorrectly or when natural language queries were submitted. Nor were suggestions for correct spelling provided by these search engines.

Thus, a system should also be able to handle spelling errors and natural language queries to help children find relevant documents using keyword search. In a comparison study between children and adults, Bilal and Kirby (2002) found that when children employed keyword search, most of their queries were single or multiple concepts, just as those of adults. However, adults employed advanced search syntax, while children did not use this.

Browsing taxonomies may also be difficult for a child because taxonomies in children's Web portals such as Kidsclick.org and Dibdabdoo.com use hierarchically structured taxonomies that may impose considerable cognitive load. Only a part of the hierarchy is displayed at any one time, and users must guess which route might eventually take them to the relevant term within the hierarchy (Large, Beheshti, Nesset & Bowler, 2006). With category search (i.e. browsing), children also have trouble finding the right category, because they have little domain-knowledge to decide which category is optimum. In addition, problems with browsing tools are mostly the result of a lack of vocabulary knowledge. Children often have difficulties understanding abstract, top-level headings, because their vocabulary knowledge is not yet sufficient to understand such terms (Hutchinson et al., 2006). We also reported this problem in our explorative study (see Section 4.5.5) in which children did not understand the abstract headings such as 'Knowing', 'Doing', and 'You'. Therefore, the formulation of headings should be adjusted to children's vocabulary knowledge, using simple, concrete search terms.

Children may not think hierarchically, as adults do, and may have trouble understanding the way in which hierarchically based categories are constructed. Knowing what their understanding of categories is can therefore be of great value in designing browsing tools. Bar-Ilan and Belous (2007) tried to understand which browsable, hierarchical subject categories children create by conducting a card sorting experiment with twelve groups of four children in the age range of 9 through 11 years. They suggested terms to the children through 61 cards. The children were free to add, delete or change terms. The researchers found that the majority of the category names used by existing directories were acceptable for the children and only a small minority of the terms caused confusion. Finally, information in browsing systems is often alphabetically displayed, requiring good alphabet skills. Many children have problems with alphabetizing and therefore have trouble finding information in such browsing systems (Borgman et al., 1995).

5.2.4 Children's search behaviour characteristics

Bilal (2000) found in her research on the use of the Yahooligans! Web Search Engine that children were chaotic in their search performance: they frequently switched between types of searching (i.e. keyword search or browsing), they often looped their keyword searches and selected hyperlinks, and they frequently backtracked. These findings suggested that children want to combine different search strategies during one search task.

Bilal and Kirby (2002) also found that children were more chaotic in their search performance than adults. In their research, Bilal and Kirby (2002) compared search behaviour between twenty-two children (aged 12 through 13) and twelve graduate students. The children made more web moves, they looped searches and hyperlinks more often, they backtracked more often, and they deviated from a designated target more often. The researchers concluded that adults adopted a 'linear or systematic' browsing style whereas most children had a 'loopy' style. They explain that this 'loopy' style can be caused by children's lower cognitive recall, because the web imposes memory overload that reduces recall during navigation. They also found that children scrolled result pages less often than adults.

We should keep in mind that most of these studies were conducted in a time that children did not make use of computers and the Internet as much as they do anno 2013. Children nowadays are much more experienced users of digital interfaces because of iPads, Facebook, online gaming, and the like, which makes it difficult to apply these research results to children's current information searching and navigation behaviour on digital interfaces. However, in our explorative study, which was conducted in 2010, many children also clearly showed a chaotic, 'loopy' navigation style (see Section 4.5.5). This result suggests that experience with the internet is not related to a less chaotic

navigation style and that this chaotic navigation style is inherent to the cognitive developmental stage that children are in.

What we have learned so far from this theoretical background is that playful interface design emerges in the genre of children's informational websites (Chapter 3). Researchers that study children's search behaviour and the problems and successes that children experience during information searching, especially discuss pragmatic issues such as query handling and comprehensibility of taxonomies (see Section 4.5.5). However, the emergence of playful interface design asks for a broader focus than pragmatic issues. Also hedonic issues of playful interface design should be studied. It is assumed that product characters can be described by two attribute groups: pragmatic and hedonic attributes. Pragmatic attributes are connected to the users' need to achieve goals (e.g. finding information on an informational website). Hedonic attributes are primarily related to the users' self. A product can be perceived by users as hedonic because it provides stimulation by its challenging character or identification by communicating personal values to relevant others (Hassenzahl, 2004). By analyzing children's emotional expressions while conducting search tasks, we found that many emotions expressed on the playful Image map website were caused by hedonic issues, such as joyfulness and tingling about the "fun stuff", but also irritation about "all the distracting animations" (see Section 4.6.3). Hedonic issues of interface design will be discussed in more detail in the next part of the theoretical background.

5.2.5 Fun and engagement

From the beginning, research on interaction with digital interfaces has been dominated by pragmatic issues such as the utility and usability of these systems (Thüring & Mahlke, (2007). This is the same for research on children's interaction with digital interfaces (Borgman et al., 1995; Bilal, 2000; Druin, 2003; Hutchinson et al., 2005). Usability, in particular, is a key concept for capturing the quality of use of digital products in which effectiveness and efficiency of system use is measured. The third component of the usability concept is 'user satisfaction'. Although this is measured using subjective judgments of users, these are mostly based on efficiency and effectiveness of interface usage.

In the field of Interaction Design for Children (IDC), there is a strong downplay in research about efficiency and task completion (Yarosh, Radu, Hunter & Rosenbaum, 2011). Instead of usability and satisfaction, that are goal related, desirability (being 'cool') has become very relevant in the community. This is not goal related. Malone (1980) pioneered the study of fun as an important aspect of software, and published guidelines designing for fun (Malone, 1984). According to his constructivist view, children acquire knowledge through experience. But for many years the study of fun in software was of marginal interest. In recent years there has been increasing interest in

fun (Read, MacFarlane & Casey, 2002). Yarosh *et al.* (2011) report that 'enjoyment' and 'fun' are the most important values in 24% of the papers presented on the yearly Conference on Interaction Design & Children. They even claim that the values enjoyment and fun are so ubiquitous in the community that they are no longer explicitly discussed, but that they have become general assumptions when designing interactive products for children.

5.2.6 User experience

Also in the general field of Human Computer Interaction (HCI) researchers argue for a broader perspective on user experience (UX) (Hassenzahl & Tractinsky, 2006; Thüring & Mahlke, 2007) which can include, - besides perceived usability - beauty, overall quality and hedonic, affective and experiential aspects of the use of technology (Van Schaik & Ling, 2008). Considering this broader perspective on UX, the reason for designing playful interfaces for children – as described in the Introduction of this chapter - becomes more clear. Playful design might have a positive effect on children's overall appraisal of a digital interface. This hypothesis is based on the idea that overall appraisal of a digital product is influenced by perception of both instrumental qualities (for example, effectiveness of a product) and non-instrumental qualities (for example, beauty of a product) as proposed by the Components of User Experience model (Thüring & Mahlke, 2007 – see Figure 2.12). Following this idea, playful design might have a positive effect on children's perception of hedonic quality because children might feel stimulated by the creative and innovative interaction style or they might be able to easily identify with the playful environment. These are the hedonic attributes of stimulation and identification that are primarily related to the users' self as described by Hassenzahl (2004). Thüring and Mahlke (2007) propose that both perception of instrumental and non-instrumental qualities have an influence on users' emotional reactions (such as subjective feelings, motor expressions or physiological reactions), which also have influence on overall appraisal of a system. For example, a system that is slow (instrumental quality) may lead to frustration (negative emotion). At the same time, this slow system might be presented via a creative interface design that may lead to enjoyment (positive emotion). Both of these emotions, if experienced, will have an influence on the overall appraisal of the system.

5.2.7 The interplay between components of user experience

What is interesting for designers of digital products to know is how the overall quality of an interactive product is formed. Evaluating interactive products is very complex because many factors influence the quality of an interactive product: usability, beauty,

overall quality, hedonic quality, and affective and experiential aspects of the use of a product.

Tractinsky, Katz and Ikar (2000), conducted an experiment to test the relationship between user's perceptions of computerized system's aesthetic beauty and usability. Perceptions were measured before and after actual use of the system. Both pre and post-use measures indicated strong correlations between perceived aesthetics and usability. Post-use usability ratings were not affected by actual usability (i.e. objective measured usability), which made Tractinsky *et al.* (2000) conclude that a product's beauty is a stronger indicator for its perceived usability than its actual usability. In other words, they claimed that "what is beautiful is usable". Tractinsky *et al.* (2000) propose the occurrence of a so-called halo-effect. The beauty of an interface overrules all other interface characteristics and therefore influences users' overall evaluation of the system.

Hassenzahl (2004) also studied the relation between perceived aesthetics and usability. He investigated the interplay between two product evaluations, beauty and goodness, and the following perceptions of product attributes: pragmatic quality (i.e. usability as perceived by the user), hedonic quality of stimulation (personal) and hedonic quality of identification (social). He found that beauty as an evaluation was related to the hedonic quality of identification (e.g. a product is perceived as professional, valuable or presentable, etc.). Hassenzahl (2004) found that goodness was more closely related to attributes of pragmatic quality (e.g. a product is perceived as simple, practical, clear, predictable, etc.), especially when participants also interacted with the product under evaluation. These results of Hassenzahl (2004) contradict the results of Tractinsky *et al.* (2000), because Tractinsky *et al.* (2000) found no significant main effect of usability on post-use ratings of usability and beauty. In contrast, Hassenzahl (2004) did find an effect of actual usability on perceptions of usability. Hassenzahl (2004) explains this contradiction in results by the fact that Tractinsky's manipulation of usability was unlikely to induce stress for the participants, which makes any impact on post-use ratings of usability unlikely.

To study how the overall quality or goodness of an interactive product is formed, Van Schaik and Ling (2008) also conducted an experiment on the interplay between components of UX. They found that all measures (i.e. evaluation of goodness, attributes of hedonic and pragmatic quality, task performance and mental effort) except evaluation of beauty, were sensitive to the manipulation of web design. Evaluation of beauty was influenced by hedonic attributes (identification and stimulation), but evaluation of goodness was influenced by both hedonic and pragmatic attributes as well as task performance and mental effort. Attributes of hedonic quality were more stable with experience (i.e. using the interactive product) than attributes of pragmatic quality. Evaluation of beauty was more stable than evaluation of goodness.

Hartmann, Sutcliffe, and DeAngeli (2008) found a link between aesthetics and usability. When users' usability experience was poor, positively perceived aesthetics could positively influence overall appraisal of a system, suggesting that "aesthetics could be an important determinant of user satisfaction and system acceptability, overcoming poor usability experience" (p. 176). Furthermore, they argued that the relative importance of aesthetics is related to the user's background and task. When the user's task is goal-oriented, then usability factors will weigh more than aesthetic considerations. When the user's task is action-oriented (the experience is more important than the goal), users choose designs based on a general impression of aesthetics and engagement.

Tuch, Roth, Hornbaek, Opwis and Bargas-Avila (2012) gave an overview of the current state of research on the aesthetics-usability relation. They made a distinction between correlative studies in which aesthetics and usability were not systematically manipulated as independent experimental factors and experimental studies in which these factors were systematically manipulated. The correlative studies showed some evidence for the relation between usability and aesthetics. However, there was only limited inference on the direction of any causality between aesthetics and usability. In the experimental studies, a pure "what is beautiful is usable" notion was only partially supported. Tuch *et al.* (2012) reported that there was also some evidence that in certain cases the relation is best described as "what is usable is beautiful" (p. 1598). Tuch *et al.* (2012) conducted an experiment on the aesthetics-usability relation and also found, under certain conditions, evidence for the relation "what is usable is beautiful". They found that the frustration of poor usability lowers ratings on perceived aesthetics.

Note that in none of the studies discussed so far were children involved. However, Hartmann *et al.* (2008) suggest that a metaphor-based interface style, such as an Image map website type (considered more aesthetically pleasing and engaging), would be better for children than a menu-based style if they were interacting with it in their leisure time. They were undecided about this if the interface was to be used in the classroom within a formal educational context. Their reason for this assumption was that a metaphor-based style would likely prove more engaging but perhaps at the expense of usability.

In a study to validate the Fun Toolkit, a tool to evaluate technology with children, Sim, MacFarlane and Read (2006) did try to relate the constructs 'fun' and 'usability'. They report that children experienced less fun when there were more usability problems. They conclude that it is not all about fun for children and that usability does matter to them.

The results of our explorative study (see Chapter 4) suggest that - when it comes to finding information - for children it is not about 'fun' at all and that it is 'usability' that matters for them. Although children expressed many positive emotions based on

hedonic issues about the playful Image map website, these positive emotions did not positively influence children's overall evaluation of the website.

We have now learned that conducting research on hedonic issues such as fun and engagement has become more important in the field of child-computer interaction in recent years. The study of pragmatic and hedonic issues of interface design with adult users from an evaluation perspective is called 'user experience' (UX). An important topic in this field of research is the interplay between components of UX, such as usability, beauty and goodness. We think that this topic is also relevant for our research on children's informational websites, especially concerning the emergence of playful interface design in this genre. Therefore, in our explorative study (Chapter 4), we explored the influence of design type on children's search performance, the emotions they displayed, and overall evaluation. However, empirical research is needed to validate the results suggested by this explorative study. Also, other research methods are needed to study children's emotions, perceptions of pragmatic and hedonic quality and overall system evaluation.

Methods used in studies with adults are mostly not suitable for studies with children. We should reflect upon whether the methods currently applied are suited for use with children. Therefore, we will now discuss the literature on the methods used in this field of research with adults and the methods used to measure hedonic components of UX with children.

5.2.8 Methods to measure the UX components

As mentioned before, research on interaction with digital interfaces has been dominated by pragmatic issues such as the utility and usability of these systems. Methods to measure usability measure effectiveness (the accuracy and completeness with which specified tasks can be conducted in a particular environment), efficiency (for example, the amount of time or the number of digital events required to reach a specified goal) and satisfaction in using the system (based on instrumental qualities of the system). The same methods to measure these factors of usability that are validated in research with adults, can be used well in research with children, as we experienced in the explorative study on children's search behaviour reported in Chapter 4.

The evaluation of subjective aesthetic preferences and emotional experiences is more difficult to measure than objective usability scores. According to Laarni (2004), this is one of the reasons why these non-instrumental qualities have played a marginal role in human-computer interaction (HCI) research. Particularly with children, measuring the subjective, non-instrumental qualities of a system, for example, with a survey method of asking children to rate product evaluations (such as fun, beauty and goodness) is very difficult, because of the risks of satisficing, children's tendency to say yes irrespective

of the question and children's tendency to indicate the highest score on the scale when scales are used to elicit opinions about software (Markopoulos et al., 2008).

Horton, Read and Sim (2011) report a study in which the reliability of children's responses on a pictoral questionnaire was tested by asking the same questions about children's technology twice, one week after the other. None of the children produced the same results for a question after one week, which proves the difficulty of using survey questions with children and the issues with the validity and reliability of questionnaire answers given by children.

Read, MacFarlane and Casey (2002) developed the Fun Toolkit to measure children's opinions of technology which reduces the mentioned risks of evaluating products with children. The Fun Toolkit consists of four tools: a Funometer, a Smileyometer, a Fun Sorter and an Again-Again table. The Toolkit has been validated in several studies with children (Read & MacFarlane, 2006; Sim, MacFarlane & Read, 2006; Read, 2008).

Visual Analogue Scales (VAS) are often used in survey studies with children. With a VAS, children can identify and visualize their answers, feelings or opinions through pictorial representations instead of textual labels (Markopoulos, et al., 2008). The Smileyometer (Read et al., 2002) is an example of a VAS. But even with such a simple question-answering style as VAS, a child still needs to understand the question, needs to recall relevant information from memory, needs to decide what response is appropriate, needs to translate this response by deciding which pictogram from the VAS is relevant, and requires to physically act to make the selection. All of these steps put high cognitive load on children's working memory that can be problematic for a child.

An often used subjective method that uses VAS to measure emotional valence and arousal is the Self-Assessment Manikin (SAM) (see Figure 5.6 in the Method section), developed by Lang (1980). The SAM is a non-verbal pictorial assessment technique that measures the pleasure, arousal, and dominance associated with a person's affective reaction to a wide variety of stimuli (Bradley & Lang, 1994) which is based on the dimensions of valence, arousal and dominance (Russell, 1980). The use of SAM with children was validated by Greenbaum, Turner, Cook and Melamed (1990).

To measure users' perceptions of three product attributes (pragmatic quality, hedonic quality – identification and hedonic quality – stimulation) and two product evaluations (beauty and goodness), Hassenzahl, Burmester and Koller (2003) developed the AttracDiff 2 questionnaire. Each of the three product attributes is represented by seven 7-point semantic differential scales (e.g. professional – amateurish) and the two product evaluations by one 7-point semantic differential scale each (e.g. good – bad). To the best of our knowledge, the AttracDiff 2 questionnaire has not yet been validated in survey studies with children.

5.3 Research questions & hypotheses

As we have learned from the literature, there is a general assumption that digital products or systems for children should be fun and engaging (Yarosh et al., 2011). Non-instrumental product attributes - like beauty and fun - are expected to have an influence on overall appraisal of a digital product, just as instrumental product attributes such as effectiveness and learnability (Thüring & Mahlke, 2007; Van Schaik & Ling, 2008).

It seems that designers of children's interactive products make their products fun and engaging by adding playful design characteristics, both in visual design and in navigation design (reported in Chapter 3). Therefore, we are interested on seeing whether the addition of playful design has a positive influence on children's perceptions of hedonic quality and on their overall appraisal of the system. We are also interested in the relation between children's perception of hedonic quality with their perception of usability and actual task performance.

In a previous explorative study (reported in Chapter 4), we did find a positive influence of playful interface design on children's perceptions of hedonic quality, measured by observing children's emotional expressions while working with informational websites in a school setting. However, most emotional expressions were based on pragmatic product attributes instead of hedonic product attributes. From this study, we concluded that playful design does not have a great influence on children's overall appraisal of informational interfaces. Usability seems much more important for children than non-pragmatic qualities, such as fun and beauty. However, that study was not set up as an experiment. Therefore, in this follow-up study, we wanted to test these initial results under controlled experimental conditions.

In this section, we will present the research questions of our experiment which will help us to answer *RQ 2*, *RQ 3* and *RQ 4* of this dissertation.

RQ 2. *What is the effect of design characteristics of informational websites for children on their search performance and on their appreciation of the search interface?*

RQ 3. *What is the relation between children's search performance and their affective responses towards the search interface?*

RQ 4. *What is the difference between browsing and searching in terms of efficiency and effectiveness of children's search task performance?*

For the research questions that can be empirically tested (i.e. research questions chapter 5-1, chapter 5-2 and chapter 5-3) a hypothesis is formulated based on the results of our explorative study. Research questions chapter 5-4 and chapter 5-5 are qualitative research questions for which no hypotheses have been formulated.

RQ – chapter 5-1: Effects of interface design *(related to RQ 2)*

a. What are the effects of different interface design types of children's informational websites on children's search performance?
   - ► Hypothesis 1a: playful 'visual' design (such as the Classical play design type) will have a positive influence on children's search performance, but playful 'navigation' design (such as the Image map design type) will be confusing and distracting for children and will therefore have a negative influence on children's search performance.

b. What are the effects of different interface design types of children's informational websites on children's evaluation of informational websites (emotions, perceptions of pragmatic and hedonic quality and product evaluations, such as beauty, goodness and fun)?
   - ► Hypothesis 1b: playful 'visual' design will have a positive influence on children's overall evaluation, based on a positive influence on emotions, perceptions of hedonic quality, and evaluation of beauty, goodness and fun.

Playful 'navigation' design will have a negative influence on children's evaluation, based on a negative influence on perception of pragmatic quality, negative emotions and a negative evaluation of goodness.

RQ – chapter 5-2: Relation between performance and evaluation *(related to RQ 3)*

What is the relation between children's performance scores (usability) and evaluation scores concerning their search behaviour on informational websites?

- ► Hypothesis 2: instrumental qualities (i.e. usability) that lead to high performance scores will have a stronger influence on children's evaluation scores than high non-instrumental qualities (such as beauty, fun and hedonic quality).

RQ – chapter 5-3: Interaction with the chosen search strategy *(related to RQ 4)*

Is there an interaction between different interface design types with the chosen search strategy (searching or browsing the main categories)?

- ► Hypothesis 3a: when browsing the main categories, children will perform better with a classic navigation design type (i.e. Classic and Classical play website) than with a playful navigation design type (i.e. Image map website).

Although, the explorative study did not provide clear directions for search engine performance, using a search engine will not be influenced by the design type of an interface. Therefore, when using the same search engine on different websites, performance by using this search engine will be the same for the different design types.

- ► Hypothesis 3b: Overall, children will perform better by using the search engine than by browsing the main categories. In the explorative study, the children performed search tasks best by using Google, which is solely based on keyword

searching instead of category browsing. The children also experienced more problems by browsing the categories than by using the search engine on the other websites in the explorative study.

RQ – chapter 5-4: Problems and successes with keyword searching and browsing
(related to RQ 4)
If any, what problems and successes do children experience when searching with a search engine or when browsing main categories?

a. Do these problems and successes relate to design characteristics of the different design types?
b. Do these problems and successes relate to the quality of the search engine?
c. Do these problems and successes relate to characteristics of the children?

RQ – chapter 5-5: Exploring research methods
Are existing methods to measure feelings and perceptions of pragmatic and hedonic quality that are used in research with adults also suited for research with children?

Before answering these research questions, we will first discuss the methods used in the experiment to be able to find answers to the research questions.

5.4 Method

5.4.1 Experimental design

The experiment used a 3 x 2 between groups design with two factors: interface design and use of the search engine. Three versions of the same website varied in aesthetics and navigation style. We did not want to have an influence on the children's natural search behaviour by telling them to search by using the search engine or to browse by using the navigation to find information. Therefore, the use of the search engine was manipulated by presenting the websites with or without a search engine. For each of the three Web design conditions, half of the children used the search engine and half of the children did not use the search engine.

Both independent variables were between-subjects; each child participant used one of the three interface designs and did (at least for one of the search tasks in the experimental session) or did not make use of the search engine at all. Outcome measures included perceptions of product attributes (pragmatic and hedonic quality), evaluations of the websites (beauty, goodness and fun), objective performance measures, subjective emotion measures and objective emotion measures.

5.4.2 Participants

There were 158 children between the ages of 10 and 12 years old that took part in the experiment (70 boys and 88 girls, see Table 5.1a), with an average age of 10.80 (SD = .65). Of these children, 67 were in the fifth grade and 91 were in sixth grade. The children were tested at four different primary schools in the Netherlands. Each child was assigned to one of the three types of website design and we checked whether the child characteristics age, gender, school, grade and experience with the internet were equally divided over the three website versions. The number of children that used the three different website versions and the number of these children that did or did not use the search engine are presented in Table 5.1b.

Table 5.1a: Age and gender of the participants

Age	**Male**	**Female**	***Total***
9	1	0	*1*
10	15	34	*49*
11	44	44	*88*
12	10	10	*20*
Total	*70*	*88*	*158*

Table 5.1b: Distribution of children over website versions and use and non-use of the search engine.

Website version	**Number of children that used the search engine**	**Number of children that did *not* use the search engine**	***Total***
Classic version	26	25	*51*
Classical play version	26	26	*52*
Image map version	25	30	*55*
Total	*77*	*81*	*158*

5.4.3 Materials and equipment

For the manipulation of the interface design, an existing online encyclopedia for children was used (i.e. Junior Winkler Prins online encyclopedia). By using a fully working, existing website, ecological validity is higher than by building a prototype website for the experiment. The disadvantage of working with an existing website is that the website comes with real life flaws. For example, the search engine on the website in our experiment did not provide query suggestions or spelling suggestions. Also, the search engine on the website was placed in the top left corner of the screen on the existing website, which is not a conventional location for the search engine (as discussed in Chapter 3). However, this location could not be adjusted for the experimental stimuli, because of technical restrictions. Besides the existing version of the website, two other fully working versions of this same website were created; each of the three versions with a different interface design. Further, for each of the three versions, again there were two

versions: one version with and one version without a search engine. In total, there were six different versions of the same website in the experiment.

The Classic version was presented with classical aesthetics and with a classical navigation style (see Figure 5.1). The Classical play version (see Figure 5.2) was presented with expressive aesthetics, but with a classical navigation style. The Image map version (see Figure 5.3) was presented with both expressive aesthetics and a playful navigation style. The three website versions were identical concerning the main categories and subcategories, the menu structure, the content and the logo, to control for effects of these factors. Also, on the deepest navigation level, all three website versions referred to the same target pages with the same layout for each of the three versions (see Figure 5.4).

The experiment ran on a laptop (Intel Core, 2,27 GHz, 4,0 GB RAM, Microsoft Windows 7 operating system) with a remote 20" monitor that the children worked on (Figure 5.5). The screen activities were recorded with Morae usability software (Techsmith) and video and audio recordings were made with a webcam. The children filled out an online questionnaire developed in PX Lab, an open source collection of Java classes and applications for running psychological experiments (Irtel, 2007). Children's electrodermal responses (physiological measure of emotional arousal) were measured with a Q Sensor (Affectiva) (Poh, Swenson, Picard, 2010).

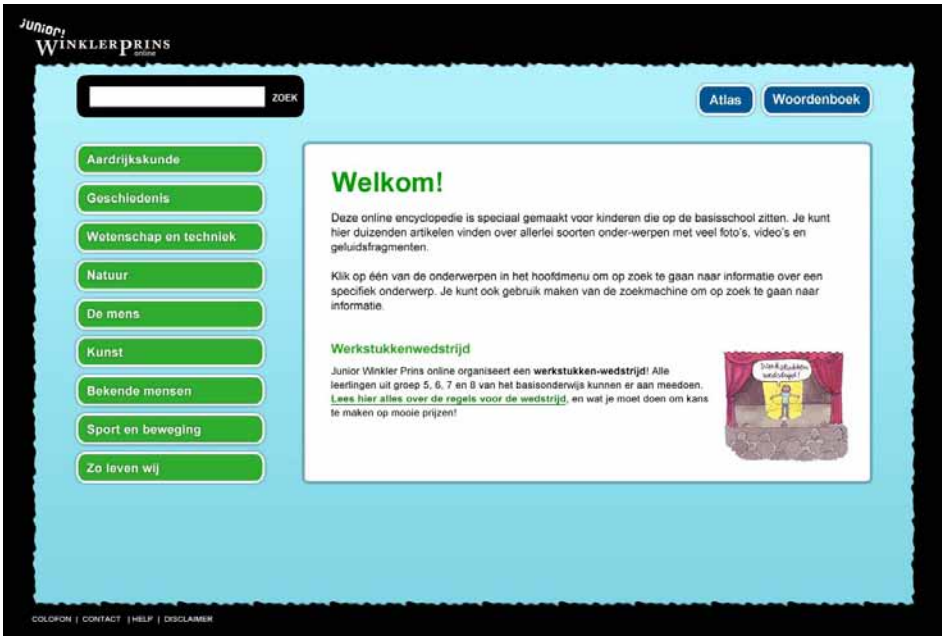


Figure 5.1: Classic website version

Figure 5.2: Classical play website version

Figure 5.3: Image map website version

5.4.4 Data collection

Measuring performance

Each child conducted the same five fact-based tasks (see Appendix D for the full task descriptions). We tested ten tasks in a pilot test with a group of 14 children. Based on

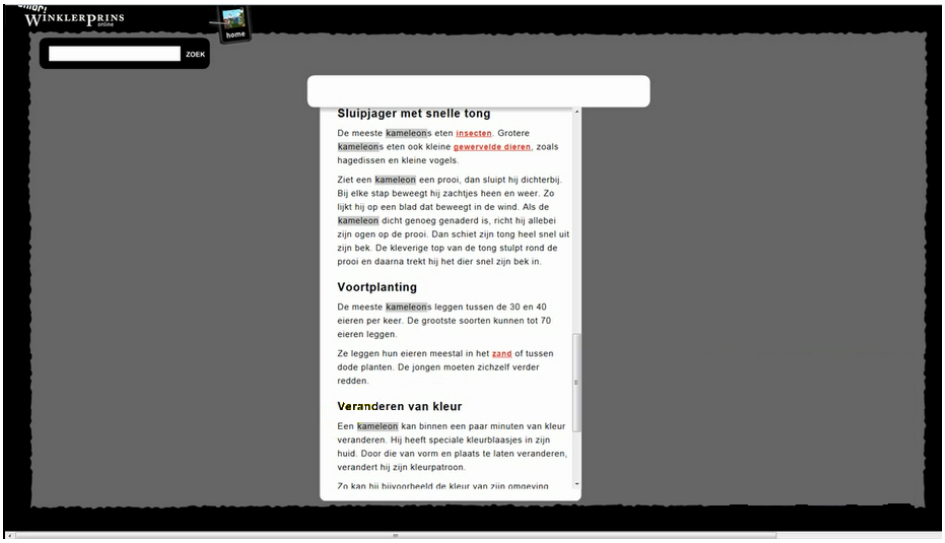


Figure 5.4: Example of a target page that is identical for all three website versions


Figure 5.5: Experimental setting

the results of this pilot test, we selected five tasks for the final experiment that varied in difficulty of conducting the task with the search engine or by browsing the categories. The task about Columbus, for example, was difficult to conduct both by using the search engine or by browsing the categories, because children had to find out which nation had discovered America 500 years before Columbus discovered America. The answer to this question was not mentioned on the content page about 'Christopher Columbus', but only on the content page about 'Discoveries' or on the content page about the 'Vikings'.

Task performance was measured by logging the amount of time and number of clicks needed to conduct the tasks. Also, per task, we analyzed whether the relevant Web page was found and if, and, if so what type of, help was offered. Recordings of the screen activities and video (see, for example, Figure 5.10) and audio recordings of the children were qualitatively analyzed, for example, to indicate what problems children experienced with any particular design characteristics.

Measuring affectivity

The children gave responses to an online questionnaire to measure their feelings. It consisted of three parts: 1) the Self-Assessment Manikin (SAM) (Lang, 1985) to measure children's valence and arousal concerning the website versions, 2) an adaptation of the AttracDiff 2 (Hassenzahl et al., 2003) to measure children's perceptions of pragmatic and hedonic quality and 3) a questionnaire to measure children's evaluations of beauty, goodness and fun by giving a report mark from 1 (= lowest score) to 10 (= highest score). Each of the items in the questionnaire was presented underneath a picture of the home page of the website version that the child had used to conduct the search tasks and that had to be evaluated by the child (for example, Figure 5.7). An added bonus to this study was that we could also evaluate whether this method was suited to measure product affectivity with children.

In the first part of our questionnaire, we used the same 5-point bipolar scale version of the SAM as used by Greenbaum *et al.* (1990) (see Figure 5.6) instead of the original 9-point bipolar scale (Lang, 1980). We decided to only measure the dimensions of valence and arousal and to leave out the dimension of dominance, as Thüring and Mahlke (2007) also did.

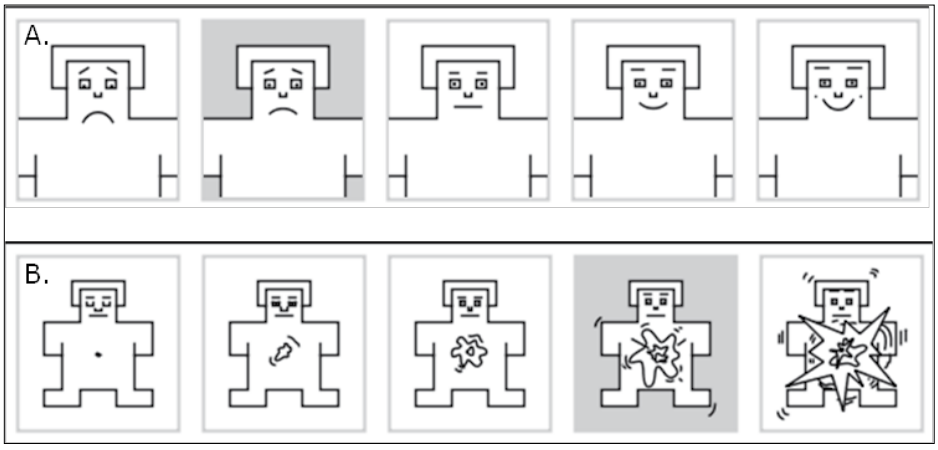


Figure 5.6: SAM 5-point bipolar scales for valence (A) and arousal (B)

Figure 5.7: Example of the bipolar picture version of the semantic differential scale 'Clear – Confusing' presented below the home page of the website version under evaluation. (The scores entered by placing the yellow cursor somewhere on the scale, were automatically recorded by the PX Lab software. The left end of the scale was scored as 1 and the right end of the scale as 5.)

For the second part of our questionnaire, we developed Visual Analogue Scales (VAS) based on the AttracDiff 2 questionnaire (Hassenzahl et al., 2003). We translated 15 of the 21 bipolar verbal anchors for the product attribute groups from the AttracDiff 2 questionnaire in bipolar picture anchors as presented in Appendix E (for example, Figure 5.7: bipolar picture anchor of the semantic differential scale 'Clear – Confusing')[7]. We used these 15 items in the pilot test and asked the children to explain the meaning of the pictures in the visual versions of the semantic differential scales. We decided to remove four of the items from the questionnaire (marked with an asterisk in Appendix E), because none of the children could give a meaning to the pictures that was close to the original meaning of the semantic differential scales. We decided to use 5-point scales instead of 7-point scales. Younger children tend to respond in an extreme manner when asked to use Likert rating scales, whereas older children are more capable of providing graded ratings in the middle of the scale. As tasks become more subjective and emotion focused, as is the case in our study, children's extreme scores, regardless

7. Six of the AttractDiff 2 items were not used in our study, because these verbal anchors were too difficult to translate in a picture that we thought could be well-interpreted by children: isolating – integrating, gaudy – classy, takes me distant from people – brings me closer to people, typical – original, conservative – innovative, commonplace – new.

of age, increase (Chambers & Johnston, 2002). Therefore, although the children in our study were between 10 and 12 years old, providing them with more than three ratings in the middle of the scale, would not add value to the rating scales. For younger children, 3-point scales would probably be most suited, because of their tendency to give extreme ratings. However, children in the age of our study (i.e. 10-12 years old) are capable of differentiating between more and less extreme ratings on a 5-point scale. We did not use bipolar picture anchors for the product evaluations of beauty, goodness and fun. Instead of that, in the final part of the questionnaire we asked the children to give report marks for the beauty, goodness and fun of the websites.

Product experience is a multi-faceted phenomenon that involves feelings, behavioural reactions, expressive reactions, and physiological reactions (Desmet & Hekkert, 2007). We measured feelings using the online questionnaire as described at the beginning of this section. We also made audio and video recordings of the children, through which behavioural and expressive reactions could be observed and analyzed. Following our qualitative analysis of verbal emotional utterances in the explorative study (see Section 4.6) we also conducted such a qualitative analysis on a smaller sample of the total number of children that participated in the experiment. We chose to analyze the emotional expressions of 25 children that used the Image map website in the experiment. We only analyzed expressions uttered with the Image map website, because in the explorative study, children uttered far more emotional expressions when using this website type than when using the other two website types. This makes observing and analyzing the expressed emotions for this website type more interesting and valuable than for the other two website types[8]. The sample of 25 participants was representative for the total group of participants in the experiment. The average age of the children in the group of 25 participants for the qualitative analysis of emotional expressions was 10.76 (average age of total sample size was 10.80). The percentage of girls in the group of 25 participants was 60% (56% in total sample) and the percentage of boys was 40% (44% in the total sample). The percentage of children in the 5th grade in the group of 25 participants was 32% (42% in total sample) and the percentage of children in the 6th grade was 68% (58% in total sample).

However, because of the risks of effects of satisficing (children's tendency to give superficial responses that generally appear reasonable or acceptable), suggestibility (the influence of the interviewer or evaluator on the children's question-and-answer process) and children's tendency to give extreme scale ratings (Markopoulos et al., 2008), we also used a more objective method to measure product experience by measuring phy-

8. Jiske Naber and Marjolein Makkinga (students of the Master Communication studies) assisted by analyzing verbal utterances of 25 children that conducted search tasks on the Image map Website.

siological reactions. We measured children's physiological emotional arousal with the Q Sensor.

The Q Sensor is a wearable, wireless biosensor that measures emotional arousal via skin conductance (SC), a form of electrodermal activity (EDA) that grows higher during states such as excitement, attention or anxiety and lower during states such as boredom or relaxation. The sensor also measures temperature and activity (Affectiva.com). Typically EDA is recorded as skin conductance by applying a direct current (with two silver electrodes) to the skin (i.e. exosomatic method). Central to this measure is the electrodermal response (EDR). The EDR constitutes a sharp rise in the SC value, followed by a slower drop in conductance. For example, a sudden loud burst of noise will result an EDR 1-2 seconds later, and this is easily visible in the raw data signal. In general, changes in SC are closely linked to activity of the sympathetic part of the autonomic nervous system. Therefore, researchers and practitioners have taken EDA measurements as further operationalization for constructs such as attention, stress, anxiety, workload, pain, and arousal (Noordzij, Scholten & Laroy-Noordzij, 2012). When children's emotional reactions are different to the three types of website design, this might be signaled by differences in the number of EDR and the total amplitude of these EDR per minute during task performance between the three websites. In Figure 5.8 an example is presented of the output of the Q Sensor, in which we can see that the device measures three physical properties at the same time: electrodermal activity, electrode temperature and acceleration.

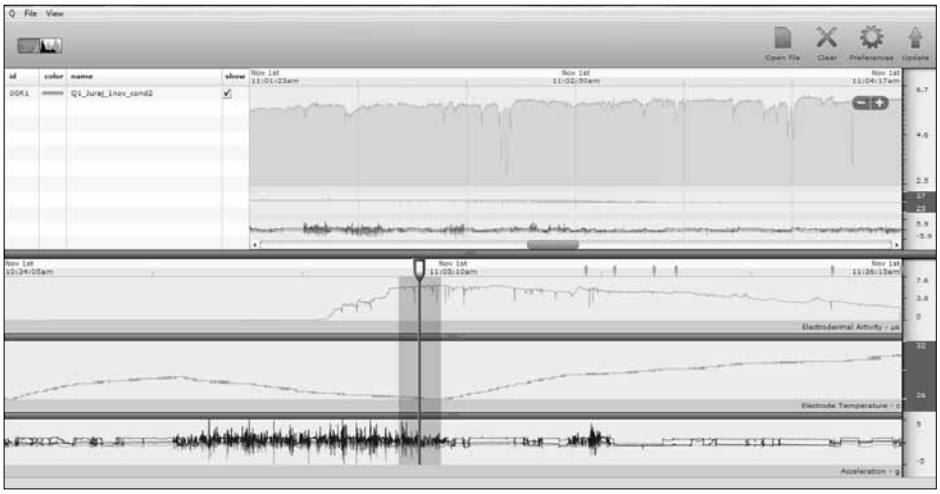


Figure 5.8: Example of the Q-sensor's output for electrodermal activity, electrode temperature and acceleration.

5.4.5 Procedure

The study was carried out in the fifth and sixth grades of four different primary schools in the Netherlands in the period of September – November 2011. Only children that could hand in a signed consent form from their parents could take part in the study. All children that took part in our study filled out a profile survey in the classroom in which we asked about their media use, such as their favorite video game, website or television show, and the amount of time they spent on the Internet, on video games, or on watching television. Half an hour before a child contributed to our study, the test instructor (i.e. the author of this dissertation) put the Q Sensor on the child's wrist. In that way, the Q Sensor could become accustomed to the child while the child stayed in his classroom. After half an hour, the test instructor came back into the classroom to take the child to the room in which the experiment was conducted. Before starting the actual task performance the child was asked to run up and down the stairs three times to activate the Q Sensor. After that, half of the group of children was asked to watch a short animation film (Disney Pixar – "For the Birds") to record a base line for the Q sensor that was the same for all 158 participants. The other half of the group of participants watched the film at the end of the session to control for a potential order effect of watching the film before or after task performance on the experimental website. This protocol of how to activate the Q sensor was formulated for this specific experiment. The protocol is not based on previous protocols, because to our knowledge the Q Sensor was not yet used for this type of research with children.

Before the actual task performance started, the test instructor explained that the child would be asked to conduct five search tasks on a website, because the designer of the website wanted to know how the children felt about the website and whether the children could easily find information. The test instructor emphasized that "the child was not being tested in this study, but that the website was being tested". The five search tasks were provided to the child in random order on separate sheets. When the child received the first search task, the test instructor started recording the screen activity and the video and audio recording of the child. The test instructor sat next to the child during the task performance and offered help when appropriate. Withholding help and social interaction can make the child feel uncomfortable (Markopoulos et al., 2008). We listed the type of help that was offered for each conducted search task. Types of help were help in choosing the right category or sub-category, help in operating the navigation tools, help with spelling or formulation of a search query, or help in finding the right information on a particular content page. The effects of interface design on children's search performance were corrected for the help that was offered to the children. After the search tasks were completed, the test instructor started the online questionnaire and asked the child to read and answer the questions presented in the questionnaire.

5.4.6 Data analyses

To measure the effects of interface design and search engine use on children's website performance and attitude towards the website, a multilevel model was constructed. Independent variables were the type of interface design (Classic, Classical play and Image map) and the use (or non-use) of the search engine. As will become clear in the Results section, it turned out that use of the search engine needed to be divided into three instead of two groups: 1) a group of children that used the search engine; 2a) a group of children for which the search engine was not provided (our initial manipulation of the use vs. non-use of the search engine) and 2b) a group of children for which the search engine was provided, that did not make use of the search engine.

Dependent variables were the percentage of children that were successful in finding the right answers to the search questions (i.e. the chance for success in our multilevel model), the time needed, and the number of clicks needed to find the right information. The model also estimates task variance (because one task can be more difficult than another task), between-children variance (because one child can be more or less skilled in searching information) and residual variance (for example, one child can have more difficulties with task 1, whereas another child can have more difficulties with task 2). By measuring these variances, we can estimate the extent to which we can generalize over tasks and children. When we do not take into account this task variance, between-children variance and residual variance, then the probability of falsely rejecting the null hypothesis is greater than 0.05 (Snijders & Bosker, 2012).

5.5 Results

5.5.1 Reliability and validity of the affective response questionnaire

Before we report the results concerning the effects of interface design and search strategies on children's search performance and affectivity towards the interfaces, we will first report the reliability and validity of the methods used.

We measured actual perceptions of hedonic and pragmatic quality using an instrument that was derived from the AttracDiff 2 questionnaire as composed by Hassenzahl *et al.* (2003). Cronbach's alphas for the two clusters of bipolar verbal anchors for the constructs of hedonic and pragmatic quality are shown in Table 5.2. A Cronbach's alpha of 0.6 is usually regarded as the lower bound of an acceptable reliability for experimental purposes. Both clusters measure the underlying constructs in a reliable way (see Table 5.2). However, the average scores and standard deviations show that most children chose the center of the scales instead of the extreme scale ratings, which makes the

reliability scores less meaningful. Apparently, children chose the safe, neutral ratings (the center) of the scales, for reasons that we will discuss further in this section.

The validity of the questionnaire was studied by a qualitative analysis of the recordings of all respondents that filled out the questionnaire.[9] In total, 2054 items were filled out by the 158 children and 151 of these children uttered a verbal or non-verbal interpretation of at least one of the 13 items from the questionnaire. These 151 children uttered 693 verbal or non-verbal interpretations of the items, which is 33.7% of the total items that were filled out in the experiment. The items from the questionnaire to measure beauty, goodness and fun were left out of this qualitative analysis, because none of the children indicated having problems with the meaning of these items. Of the 693 utterances in total, 224 utterances represented correct interpretations of the items (as intended by Hassenzahl et al., 2003) and 330 utterances represented incorrect interpretations of the items. For example, about the emotional arousal SAM-scale, many children gave the following incorrect interpretation: "I think this is about how easy or difficult the website is". Another 84 utterances were verbal indications of incomprehension of the meaning of the items (e.g. "I don't get it.") and 33 utterances were non-verbal indications of incomprehension (e.g. by frowning). While filling out the questionnaire, the test instructor helped the children when they asked for help, as discussed in the Method section. Most help was offered by asking a counter-question, for example: "What do you think the picture stands for?" Although, only a third of the items represent children's interpretations of these items, we think that these interpretations can be related to the entire set of items that were filled out in the experiment. The interpretations of the items will be discussed in the next section.

Table 5.2: Cronbach's alphas of the constructs of hedonic and pragmatic quality (using 5-point Likert scales from 1 to 5).

Cluster	Scale	$N_{\text{semantic differentials}}$	Cronbach's alpha
Hedonic quality	5-point	5	.63
Pragmatic quality	5-point	6	.72
All semantic differentials	5-point	11	.80

Interpretations of the SAM-scales for emotional valence and arousal

Most children (138 of the 158 children) did not utter verbal or non-verbal interpretations of the valence scale (see Figure 5.6A). Nor did the children ask many questions concerning this item. Most interpretations of this item were correct and were related to positive and negative emotional feelings, such as "The website is kind of fun". Apparently, children understood the meaning of this valence scale with a more or less smiling manikin, which is related to the Smileyometer of the Fun Toolkit (Read et al., 2002).

9. Aafke Ariaans (MA student) assisted by analyzing the reliability and validity of the data.

Therefore, we decided that the valence scale was a valid method to measure children's emotional valence towards a website in our study.

The SAM scale for emotional arousal (see Figure 5.6B) caused many more problems than the scale for emotional valence. In total, 85 children uttered an incorrect interpretation or an indication of incomprehension concerning this scale and 32 children uttered a correct interpretation. Most conspicuous interpretations of the SAM-scale of arousal were: "I don't get it! Why is his belly exploding?" (while he points at the picture for highest arousal), or "I haven't got a clue. It looks like a fried egg on his belly, or something like that". Obviously, for most children the meaning of the SAM-scale for arousal was not clear at all and therefore, in our experiment, the arousal scale has proved to be an invalid method to measure children's emotional arousal towards a website.

Interpretations of the bipolar picture anchors for pragmatic quality

The picture anchors for the semantic differential 'technical – human' were most problematic of all pragmatic semantic differential items for the children as can be seen in Table 5.3. Only one child gave the correct interpretation of the picture anchors. Many children associated the hearts and flowers with hedonic concepts of 'love' and 'fun', while it was intended as a pragmatic concept. Also, making a direct translation of these pictures to working with a website was very difficult for the children. The reliability of the pragmatic scale items increased to .72 if this problematic item 'technical – human' was deleted from the list.

Almost half of the interpretations uttered on the picture anchors for the concept 'complicated-simple' were incorrect. Many children (27) asked for help interpreting these pictures, which also stresses the difficulty of these pictures. Some children gave a literal meaning to the pictures, such as 'neat or scratchy lines' and associated this with hedonic concepts. They could not make a translation to the pragmatic concepts of 'simple' and 'complicated'.

The same problem of literal translation of the pictures was the case with the picture anchors for 'impractical – practical'. Children asked what "tripping over a stone has to do with searching on a website?" The interpretation 'easy – difficult' was often given to these picture anchors.

Although many children seemed to understand the meaning of the picture anchors for clear – confusing, they interpreted the meaning quite literally, by saying that the main and submenus offered many options to choose from. However, they did not give their opinion about whether these options were clear or confusing.

Most children gave a correct interpretation of the picture anchors for 'cumbersome – direct'. However, help was often asked from the test instructor and 25 children received help by giving them a counter-question or by explaining the meaning of the picture anchors. Many children gave the correct interpretation "whether you can find it directly

or with a detour". However, most of them based their answer on their own performance instead of on the directness or cumbersomeness of the website.

The pragmatic item that was least problematic for the children was the semantic differential 'unruly – manageable'. Most interpretations given for this item were correct and little help was offered by the test instructor. The reason for this is that a literal translation of the picture anchors can be directly related to a website, because a website can look neat or unruly.

Table 5.3: Frequency table of uttered interpretations of the pragmatic semantic differential items (N = absolute number of children that gave the interpretations).

	Technical-human (N=158)	**Complicated – simple** (N=158)	**Impractical – practical** (N=158)	**Confusing – clear** (N=158)	**Cumbersome – direct** (N=158)	**Unruly – manage-able** (N=158)
Incorrect interpreta-tions	38	28	15	17	11	8
Correct interpreta-tions	1	27	20	24	35	34
Verbal indication of incomprehension	18	2	2	2	1	0
Non-verbal indication of incomprehension	8	1	1	3	2	1
No utterances	90	98	118	112	106	118
Missing values	3	2	2	0	3	0

Interpretations of the bipolar picture anchors for perceived hedonic quality

The hedonic items were even more problematic than the pragmatic items (see Table 5.4), because the interpretations of these concepts were often even more difficult to translate to the use of a website. The picture anchors 'easy – challenging' were given the most incorrect interpretations by the children, which was 'easy – difficult'. To strictly test the validity of the questionnaire, we decided to score these interpretations as incorrect, because 'easy – difficult' is a pragmatic concept as opposed to the hedonic concept 'easy – challenging'. The reliability of the hedonic scale items increased to .73 if this problematic item 'easy – challenging' was deleted from the list.

The picture anchors for 'cheap – valuable' caused a lot of problems for the children, because first of all, they did not understand the meaning of the paper hat versus the crown. When the test instructor asked: "What do you think is the difference between the two hats?", children gave the interpretation of 'poor versus rich'. However, they did not understand how a website could be 'poor or rich'.

The children did not understand the picture anchors for 'amateurish – professional' at all. They often thought it was about 'an electric drill versus a hammer' and they could not relate these concepts to a website.

The picture anchors for 'presentable – unpresentable' evoked a lot of questions for the test instructor. When children gave the interpretation of an 'old or new present',

or maybe even a 'beautiful or ugly present', they could not relate this 'present' to the website.

Finally, the picture anchors for 'lame – exciting' were also problematic for the children. Again, the children had trouble understanding the meaning of the two types of 'cycling' to 'lame and exciting'. When the test instructor gave help with this first step, most of the time the children gave their opinion about whether searching for information is lame or exciting and not whether the website was lame or exciting.

Table 5.4: Frequency table of uttered interpretations of the hedonic semantic differential items (N = absolute number of children that gave the interpretations).

	Easy – challenging (N=158)	**Cheap – valuable** (N=158)	**Amateurish – professional** (N=158)	**Unpresentable – presentable** (N=158)	**Lame – exiting** (N=158)
Incorrect interpretations	51	34	33	24	8
Correct interpretations	2	6	3	15	15
Verbal indication of incomprehension	7	16	5	2	2
Non-verbal indication of incomprehension	2	4	3	3	2
No utterances	95	97	113	113	127
Missing values	1	1	1	1	4

Distinction between perceptions of pragmatic and hedonic quality

We also measured the reliability of all semantic differentials together and it turned out that Cronbach's alpha for all semantic differentials was .80 (see Table 5.2). It seems that the children perceived all items as the same construct, for example as the construct 'good or bad'. One child said: "It's all a bit the same to me" and often children asked for confirmation: "So this means good and this means bad, right?" In other words, the children did not make a distinction between hedonic and pragmatic constructs or between fun and usability.

And although the children did not indicate having problems with the report marks for beauty, goodness and fun, it can be expected that these concepts were evaluated as the same construct. This was also reported by Read *et al.* (2002) concerning the Fun-Sorter in which children needed to sort products by concepts as 'worked the best', 'liked the most', 'most fun' and 'easiest to use'. They say: "This was quite difficult for the children with the result that some constructs turned out to be quite similar."

Conclusion validity questionnaire items

From the qualitative analysis of children's interpretations of the questionnaire items, we can conclude that, although the construct reliability of the questionnaire items was high, the questionnaire was not a valid method to measure children's perceptions of pragmatic and hedonic quality. The content validity cannot be guaranteed, because

often children gave another explanation to the items than was intended by the designers of the questionnaire. Also, construct validity cannot be guaranteed, because items that should measure pragmatic quality, were associated with hedonic quality and vice versa. Besides that, it seems that children do not make a distinction between pragmatic and hedonic constructs at all. Also the SAM-scale for arousal has not proven to be a valid method to measure emotional arousal with children.

The most important problem that children experience with the questionnaire is that they have to interpret the picture anchors and to relate their meaning to using a website. As children tend to take the picture anchors very literally, translating them to a more abstract concept is very difficult for them. The cognitive load of this task on children's working memory is too heavy for children to cope with. Furthermore, children tend to relate the picture anchors to their own performance or preferences instead of to the website under evaluation. Finally, satisficing is a relevant problem in our experiment. Children are prone to satisficing as they find survey participation difficult (Markopoulos et al., 2008), as was the case in our study. From the recordings, we saw that children tend to ask questions about the first four or five items in the questionnaire, but after that fill out the questionnaire very quickly. It is clear that the children gave more or less superficial responses that generally seemed acceptable, but without going through all the steps involved in the question-answer process.

Based on the validity analyses of the affective survey question, unfortunately, we can only work with a few items for further qualitative analysis of the affective data. We will work with the results from the SAM-scale for emotional valence, with the semantic differential scale for unruly-manageable and with the report marks for the product evaluations beauty, goodness and fun.

5.5.2 Effects of interface design and search strategy used

We will first report the pragmatic effects on children's search performance of differences in interface design (i.e. Classic, Classical play and Image map) and the search strategy used (i.e. keyword searching or browsing). After that, we will report the effects on children's emotional feelings, perceptions of hedonic quality and product evaluations of interface design and the search strategy used.

What are the effects on task performance of playful interface design and use of the search engine?

To establish whether there is a difference between task performance on the three different versions of the website and between use and non-use of the search engine, the mean percentages for success were compared (see Table 5.5). Non-use of the search engine occurred in two situations; 1) when the search engine was not offered, children logically could not use the search engine and 2) when the search engine was offered,

some children did not use it to search for information. We also compared the mean percentages of finding the right answer when help was offered and when no help was offered by the test instructor and we found significant effects of provided help (see * in Table 5.5). The following data analyses for the mean percentages for success are therefore corrected for help (mean percentages for help are grey colored in Tables 5.5, 5.6 & 5.7).

No main effect on task performance was observed for the design type of the websites: there was no significant difference for the percentages of success in finding the right information between the three website versions (χ^2 = 1.02; df = 2; p = 0.31). Hypothesis 1a is therefore rejected for success in finding the right information.

However, a main effect was found for the use of the search engine (χ^2= 43.19; df = 2; p <.001): the percentage of success was much larger when the search engine was used than when the search engine was not offered (χ^2 = 27.33; df = 1; p <.001) and when the search engine was not used (χ^2 = 40.63; df = 1; p < .001). There was no significant difference between the percentage of success when the search engine was not offered and when the search engine was not used (χ^2 = 0.41; df = 1; p = .52). Hypothesis 3b is therefore accepted for success in finding the right information.

No interaction effect was found for the use of the search engine and the three website versions ($\chi^2 \leq 3.04$; df = 2; $p \geq 0.080$). In other words, the differences between use and non-use of the search engine for success were the same for the three website versions. Hypothesis 3a is therefore also accepted for success in finding the right information.

Table 5.5: Percentages of success (logits in parentheses) in using the different versions of the website for use and non-use of the search engine and for help provided yes (grey colored) or no.

Website version	*Search engine used*		*Search engine not provided*		*Search engine provided, but not used*	
	No help	*Help*	*No help*	*Help*	*No help*	*Help*
Classic	.82 (1.53)	.83 (1.61)	.63 (.52)	.68 (.75)	.57 (.30)	.67 (.69)
Classical play	.84 (1.62)	.63 (.51)*	.63 (.54)	.58 (.32)	.49 (-.03)	.56 (.22)
Image map	.92 (2.42)	.76 (1.18)*	.52 (.10)	.25 (-1.10)	.61 (.44)	.54 (.17)

Note. In all cases, a higher mean score represents a higher percentage for success in finding the right information for the search task. The answers for the binomial success score (1= successful, 0= unsuccessful) are given in Logits that are used for the data analysis (in parentheses).

*There is a significant effect of provided help on the mean percentage of success. The percentage of success in finding the right information was significantly lower for the children that used the search engine and received help from the test instructor for both the Classical play website and the Image map website (t ≥.2.01; p ≤ .04).

What are the effects of playful interface design and use of the search engine on time and clicks needed to conduct the tasks?

We also compared the mean amount of time and number of clicks children needed to conduct the tasks between the three website versions and between the use and non-use

of the search engine (see Table 5.6 and 5.7). As we also found significant effects of the help provided for these factors (see * in Table 5.6 and 5.7), we corrected the data for provided help.

Table 5.6: Mean time needed in seconds (ln in parentheses) using the different versions of the website, for use and non-use of the search engine and for help provided yes (grey colored) or no.

Website version	*Search engine used*		*Search engine not provided*		*Search engine provided, but not used*	
	No help	*Help*	*No help*	*Help*	*No help*	*Help*
Classic	139.9 (4.94)	247.4 (5.51)*	168.3 (5.13)	264.4 (5.58)*	157.9 (5.06)	300.1 (5.70)*
Classical play	149.7 (5.01)	257.0 (5.55)*	141.6 (4.95)	283.9 (5.65)*	142.7 (4.96)	221.2 (5.40)*
Image map	112.8 (4.73)	236.0 (5.46)*	170.4 (5.14)	164.1 (5.10)	148.5 (5.00)	279.7 (5.63)*

Note. The distribution of the raw data for time was not comparable to the normal distribution. Therefore, we took the natural log of the search times that did show a normal distribution.
* There is a significant effect of provided help on the mean amount of time needed. The amount of time needed to conduct the tasks was significantly higher for children that received help from the test instructor ($t \geq 3.73$; $p \leq .001$).

Table 5.7: Mean number of clicks (ln in parentheses) using the different versions of the website, for use and non-use of the search engine and for help provided yes (grey colored) or no.

Website version	*Search engine used*		*Search engine not provided*		*Search engine provided, but not used*	
	No help	*Help*	*No help*	*Help*	*No help*	*Help*
Classic	4.2 (1.44)	7.4 (2.00)*	9.8 (2.29)	17.0 (2.83)*	7.7 (2.04)	17.6 (2.87)*
Classical play	4.7 (1.55)	8.1 (2.09)*	7.7 (2.04)	14.9 (2.70)*	7.9 (2.07)	8.5 (2.14)*
Image map	3.3 (1.20)	7.2 (1.97)*	9.0 (2.20)	8.3 (2.12)	6.3 (1.85)	14.4 (2.67)*

Note. The distribution of the raw data for number of clicks was not comparable to the normal distribution. Therefore, we took the natural log of the number of clicks that did show a normal distribution.
* There is a significant effect of provided help on the mean number of clicks. The number of clicks needed to conduct the tasks was significantly higher for children who received help from the test instructor ($t \geq 2.85$; $p \leq .004$).

No main effect was observed for the design of the website versions: there was no significant difference in time needed to conduct the tasks between the three website versions ($\chi^2 = 2.00$; df = 2; p = 0.37). Hypothesis 1a is therefore rejected for the time needed. However, a main effect was found for the use of the search engine ($\chi^2 = 8.10$; df = 2; p =.017): less time was needed when the search engine was used than when the search engine was not offered ($\chi^2 = 6.88$; df = 1; p =.009) and when the search engine was not used ($\chi^2 = 4.27$; df = 1; p < .039). There was no significant difference between the time needed when the search engine was not offered and when the search engine was not used ($\chi^2 = 0.88$; df = 1; p = .35). Hypothesis 3b is therefore accepted for the time needed.

No interaction effect was found for the use of the search engine and the three website versions ($\chi^2 \leq 2.97$; df = 1; $p \geq 0.16$). In other words, the differences between use and

non-use of the search engine for the time needed to conduct the search tasks are the same for the three website versions. Hypothesis 3a is therefore also accepted for time needed.

No main effect on clicks was observed for the design of the website versions: there is no significant difference in the number of clicks between the three website versions (χ^2 = 4.08; df = 2; p = 0.13). Hypothesis 1a is therefore rejected for the number of clicks. However, a main effect was found for the use of the search engine (χ^2 = 257.56; df = 2; p < .001): the number of clicks needed when the search engine was used was lower than when the search engine was not offered (χ^2= 80.27; df = 1; p < .001) and when the search engine was not used (χ^2 = 66.52; df = 1; p < .001). Hypothesis 3b is therefore accepted for the number of clicks.

The number of clicks was also significantly higher when the search engine was not offered than when the search engine was not used (χ^2 = 212.78; df = 1; p < .001). A reason for this could be that children in the condition without a search engine were normally used to working with a search engine and therefore, were less experienced and needed more clicks to find the information using the navigation structure.

No interaction effect was found for the use of the search engine and the three website versions ($\chi^2 \leq 3.77$; df = 1; p = 0.052). In other words, the differences between use and non-use of the search engine for the clicks needed to conduct the search tasks were the same for the three website versions. Hypothesis 3a is therefore accepted for the number of clicks.

In conclusion, children who used the search engine instead of browsing the categories were more successful in finding the right information and they needed less time and fewer clicks. Therefore, hypothesis 3b is accepted. There were no significant differences for task performance (i.e. success, time and clicks) between the three website versions. Apparently, interface design of the search environments is a lesser determinant for task performance than the search strategy (i.e. searching or browsing). Therefore, hypothesis 1a is rejected.

The differences that we found between the conditions (independently of the used search method; searching or browsing) cannot be assigned to our manipulations of the interface design. These differences should be assigned to the differences between the children that participated in our experiment. In other words, differences in child characteristics, such as their information skills, domain knowledge, operational skills, etcetera, caused more variance in the children's search performance than variation in interface design.

What is the effect of interface design on emotional valence and on the evaluation of beauty, goodness & fun

To test whether there is an effect of design on children's affective responses, we computed both between groups and within groups analyses of variance. In that way, we took into account the "effect variation" and the "individual variation". Individual variation is the variation within condition differences called "error", because we cannot explain the fact that children who were in the same two conditions - who were all treated the same two ways – had different scores. In this way, we also took into account the "subject variation", which is the variation due to subject variability. For these tests on children's affective responses, we could only use a limited set of variables that proved to be valid, as presented in Table 5.8. The data of the other subjective variables is presented in Appendix F.

There were significant differences for children's emotional valence and their evaluation of goodness between the three website versions. The children judged their feeling with the Classical play website as more positive than with the other two websites ($F_{2,155} = 3.28$; $p = .040$)[10]. This result supports acceptation of hypothesis 1b. Still, all scores were between 4 and 5 on a scale from 1 to 5, so the children were very positive about their feelings with all three types of website design.

The children evaluated the Image map website as least good ($F_{2,155} = 3.45$; $p = .034$) of the three websites. This result also supports acceptation of hypothesis 1b. We do not use the word 'worst', because with the 'least good' Image map website, the children evaluated the goodness of the Image map websites with an 8.2 on a scale from 1 to 10, which is still a very high score (see Table 5.8). There were no significant differences between the children's perceptions of hedonic quality (i.e. unruly-manageable) and between their evaluations of fun and beauty of the three websites.

Table 5.8: Mean scores for emotional valence (5-point scale), the semantic differential 'unruly-manageable' (5-point scale) and the product evaluations for beauty, goodness and fun (report mark from 1-10) for the different versions of the website (SD in parentheses).

Website version	*Valence*	*Unruly-manageable*	*Goodness*	*Fun*	*Beauty*
Classic (N = 51)	4.12 (.77)	4.25 (.95)	8.66 (1.31)	8.29 (1.56)	8.37 (1.42)
Classical play (N = 52)	4.38 (.66)*	4.10 (.79)	8.71 (.84)	8.50 (1.19)	8.86 (1.05)
Image map (N = 55)	4.05 (.68)	3.90 (.91)	8.22 (1.04)*	8.13 (1.10)	8.42 (1.03)

* There is a significant effect of the type of design on the affective responses.

10. The degrees of freedom used to assess the F-ratio are the degrees of freedom for the effect of the model and the degrees of freedom for the residual of the model.

What is the effect of playful interface design on physiological measurements of emotional arousal

We also used a more objective method to measure children's feelings towards the three website types by measuring electrodermal activity (EDA). From the EDA-data, we computed the number of electrodermal Responses (EDR) per minute and the total amplitude of the EDR per minute for the period of watching the film and the period of the actual task performance on the experimental website. Watching the film served as a benchmark, because this was the same over all three conditions. To measure whether there was a difference in EDA between conditions, we computed the difference in number of EDR per minute between the task performance and watching the film for the three conditions. Unfortunately, the EDR data (both number of EDR per minute during the task performance and while watching the film, and the difference between these two variables) were not normally distributed. Therefore, we could not use a parametric test to compute the difference in EDA between conditions and we were constrained to using a non-parametric test.

First of all, we did not measure any EDA during the actual task performance with 47 children. We did not measure any EDA with 64 children while they watched the film. Apparently, for many children, this type of scholarly tasks on a computer did not activate any electrodermal activity at all. Second, using a Wilcoxon signed rank test, there was no significant difference in EDA between task performance and watching the film.

According to an independent samples Kruskal-Wallis test, there was no significant difference in EDA between the three conditions. This result does not support acceptation of hypothesis 1b. There was, however, a significant order effect in our study. According to an independent samples Kruskal-Wallis test, there was a significant difference in EDA between children that performed the search tasks before or after watching the film. Children that performed the search tasks before they watched the film showed a higher number of EDR per minute during the task performance than children that performed the search tasks after they first watched the film. A regular univariate analysis of variance also showed no significant differences in EDA between the three conditions and it also did show a significant order effect of conducting the search tasks before or after they watched the film ($F(_{1,135}) = 23.03$, $p < .05$). Apparently, children were less aroused by the use of the search interface when they performed the search tasks after they saw the film, than when they performed the tasks before they saw the film. This order effect proved the working of the Q-sensor as a valid method to measure EDA with the children in the experiment.

What is the effect of playful interface design on verbal emotional utterances?

Our quantitative research methods did not provide enough proof to support acceptation of hypothesis 2 that instrumental qualities (i.e. usability) have a stronger influence on

children's evaluation scores than high non-instrumental qualities (such as beauty, fun and hedonic quality). Therefore, following our qualitative analysis of verbal emotional utterances in the explorative study (see Section 4.6) we also conducted such a qualitative analysis on a smaller sample of the total number of children that participated in the experiment. The selection of this smaller group of children is discussed in Section 5.4.4 in more detail.

Of the children that used the Image map website, 25 video recordings were studied to collect the verbal emotional utterances about the pragmatic and hedonic quality of the website. In total, 58 utterances were made. Of these utterances, 55 were related to pragmatic quality and 3 utterances were related to hedonic quality of the Image map website (see Table 5.9).

As can be seen in Table 5.9, most verbal utterances were negative. These utterances were mostly related to the difficulty and inaccessibility of the website for finding relevant information, such as "I can't find it at all" or "I find it difficult to search on". There were only a few positive utterances, such as "This is not really difficult or anything" and "I found it quite fast". Verbal hedonic utterances were very rare and were more related to pictures on the website than to the design of the website itself, such as a child that said: "What an ugly man" (about Columbus). Apparently, children tended to say more about their search activities than about the design of the website and verbal utterances were almost exclusively related to their perception of the websites' pragmatic (instrumental) quality. This qualitative result supports acceptation of hypothesis 2 that instrumental qualities (i.e. usability) have a stronger influence on children's evaluation scores than high non-instrumental qualities (such as beauty, fun and hedonic quality).

Table 5.9: Number of positive and negative vocal utterances related to pragmatic and hedonic quality of the Image map website.

Quality	*Number of utterances*	
	Positive	*Negative*
Pragmatic	7 (12.1%)	48 (82.8%)
Hedonic	0 (0%)	3 (5.2%)

5.5.3 Relation between factors

What is the relation between children's performance on the websites and their attitude towards these websites?

To answer this research question, we tested the difference for children's affective responses between tasks that were completed successfully and tasks that were not completed successfully. There was a significant difference between the successful and unsuccessful tasks for children's ratings of emotional valence ($F(1) = 5.07$; $p = .025$), although the difference between the ratings seems small in Figure 5.9. Children expressed more posi-

tive feelings for tasks conducted on all three website versions when the tasks were completed successfully than when the tasks were completed unsuccessfully. The significant difference between valence scores for the three website types (as described in Section 5.5.2) can also be seen clearly in Figure 5.9. There was no significant interaction effect between the website version and whether the tasks were completed successfully or not for the ratings of emotional valence. Furthermore, there were no significant differences for children's product evaluations (beauty, goodness and fun) between tasks that were completed successfully and tasks that were not completed successfully.

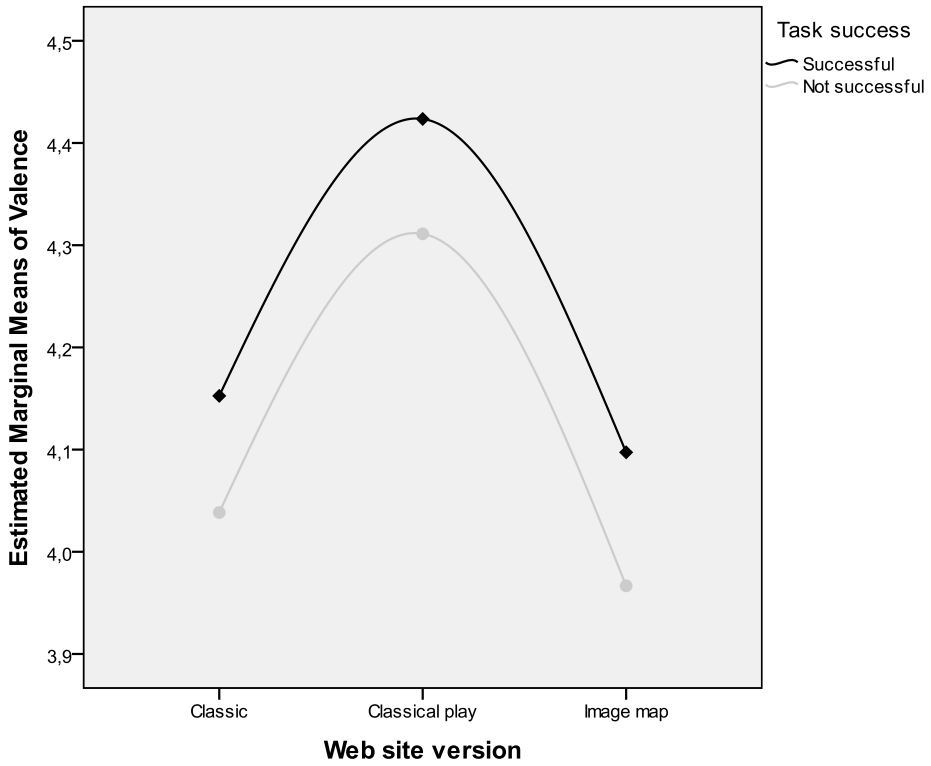


Figure 5.9: Estimated marginal means of valence (0 = negative feeling; 5 = positive feeling) plotted for the three website versions with successful and unsuccessful tasks presented on separate lines

5.5.4 Further diagnosis of children's search behaviour

Searching versus browsing in our experiment

It is not clear from previous research results whether children in general use more searching or browsing strategies. However, most of the results suggest that children prefer to use search engines, but that they are more successful with browsing categories.

Even recently, Meloncon *et al.* (2010) recommended not to include a search engine on a children's website, because children have not yet fully developed the intellectual ability necessary to generate relevant search terms (Druin et al., 2009). In addition, giving children the option to search would undermine the process of having them read through the information and explore the website (Meloncon et al., 2010). In other words, the search engine is seen as a distraction on a children's website.

In our data searching scores better than browsing the main categories for almost all components of UX that we measured. Children were more successful in finding the right content page when they used the search engine. They also needed significantly less time and fewer clicks to find the information when they used the search engine than when they browsed the main categories.

These results are in contrast to results from previous studies on children's search behaviour (Borgman et al., 1995; Schacter et al., 1998; Bilal, 2000). This difference might be caused by a better functioning search engine in our study, although we think that is hard to believe, because the search engine in our study did not work as well as Google, for example. The search engine in our study did not provide query suggestions or spelling corrections while typing a search query. Also, most natural language queries could not be processed by the search engine.

The main reason for the results in our study in comparison to previous research, is most likely caused by children's increased experience and skills in using a search engine with respect to a few years ago. Children now are members of the "Google generation" (Rowlands, 2008) and more familiar with the use of search engines. In the Netherlands in 2008, almost 80% of children aged 8-12 used the Google search engine to find information on the Internet (Pijpers et al., 2008).

In our experiment, initially 115 children were offered a website version with a search engine and 43 children without a search engine (see Table 5.10). This inequality was caused by the fact that many children that were offered a search engine, did not use the search engine at all. We offered children a version with a search engine as often as was necessary, to reach an equal number of children that used the search engine and children that browsed the main categories. In our research, we found it very important to compare searching with a search engine with browsing the main categories.

Therefore, we aimed at an equal distribution of children that did and did not use the search engine. We kept offering a website version with a search engine until the distribution between use and non-use of the search engine was equal over the three website versions (see Table 5.10). In total, 77 of the 115 children that were offered a search engine (67%), used the search engine for at least one of the five search tasks. The children that used a search engine did not automatically use the search engine for all five tasks. Therefore, the percentage of search engine use per task that was conducted on a website version with a search engine was lower. In total, 294 tasks were

conducted by using a search engine, which is 51% of all tasks that were conducted on a website version with a search engine.

As can be seen in Table 5.10, the search engine on the Image map was the least inviting to use for the children, because 25 of the 50 children that were offered a search engine on the Image map website, did not use it at all. The non-conventional visual design of the Image map version might have distracted the children's attention from the search engine as can be seen in Figure 5.3.

Another explanation for the low percentage of search engine use when a search engine was offered, can be children's infrequent use of search engines. We asked children how much time they spend using the Internet and most children used the internet less than one hour on both week days and weekend days. When we asked what type of activities they conducted on the Internet, then the activities can be ranked from most to least frequently conducted by most children as follows:

1. Hyves (a Dutch social network)
2. Playing games
3. Listening to music / watching video clips
4. Watching you tube movies
5. Searching information for myself
6. Searching information for school work

Internet activities for which a search engine is required (by searching information for themselves or for school work) were conducted least frequently according to the children's self-reports. The children probably did not use a search engine for the Internet activities that were conducted most frequently, such as using Hyves or playing games. The fact that children did not use search engines frequently is a plausible cause for the fact that many children did not use a search engine in our study, even though a search engine was provided in many of these cases.

Table 5.10: Number of children that were offered a search engine and number of children that used the search engine.

Website version *(N=158)*	**Presence of the search engine**		**Use of the search engine**	
Classic version	Present:	32	Used:	26
			Not used:	6
	Not present:	19		
Classical play version	Present:	33	Used:	26
			Not used:	7
	Not present:	19		
Image map version	Present:	50	Used:	25
			Not used:	**25**
	Not present:	5		

Search engine strategies

In this section, we will analyze children's strategies and skills using a search engine. What makes keyword searching in our experiment successful? Do children experience the same problems or not with keyword searching as reported in previous research? We will describe whether children are able to formulate a query and to select a search result that is coherent with the search task. For this qualitative study of submitted queries and selected search results, we have analyzed the submitted queries of the children that used the search engine on two of the three website types: the Classical play website (26 children) and the Image map website (25 children).

For the qualitative data analysis of search engine use, we analyzed 190 search tasks conducted with the search engine[11]. These 190 search tasks consisted of 322 search attempts to formulate a query. Most search tasks (N = 117) consisted of one query attempt, but some children required more attempts to formulate a search query within one search task (40 * 2 attempts, 20 * 3 attempts, 7 * 4 attempts, 3 * 5 attempts, 2 * 6 attempts and 1 * 10 attempts).

Quality of query attempts

Most queries consisted of one word (117 query attempts) or more than one word (165 query attempts) and 40 queries consisted of a whole sentence. Natural language querying was not as frequently applied by the children in our study as reported in previous studies (Marchionini, 1989; Druin, et al., 2009).

Most query attempts did not require any query formulation or spelling help from the test instructor (87% of all query attempts). The children had fewer problems with query formulation and spelling than expected from the literature (Borgman et al., 1995; Hutchinson et al., 2005). This was also shown by the fact that 79.5% of the query attempts were spelled correctly. An important reason was that the children in our study received the search tasks written on a task sheet, so they could use the correct spelling of words from the written task sheet. However, spelling mistakes were still made, although children could read the correct spelling from the task sheets.

Correct spelling does not automatically lead to finding the right content page, because almost half of the correctly spelled queries (44.5%) did not lead to a successful search result. Of course, the success of a search query also depends on the relevancy of the search query and the quality of the search engine. One way to evaluate the quality of the search queries is to compute their latent semantic analysis (LSA) scores in comparison to the search task, following the example of Kitajima, Blackmon, and Polson (2000) that used LSA to simulate Web navigation. LSA is a technique in natural language processing, in particular in vectorial semantics, of analyzing semantic coherence between a set of

11. Yvonne Joosten (MA student) assisted by conducting qualitative analyses of search engine use.

documents and the terms they contain by producing a set of concepts related to the documents and terms. LSA assumes that words that are close in meaning will occur in similar pieces of text. Values close to 1 represent high semantic coherence while values close to 0 represent low semantic coherence. By calculating the LSA scores of the search queries in comparison to the search task, we can evaluate the match between the query and the search task and, therefore, the quality of the query. For the calculation of LSA-scores, spelling mistakes were corrected. We did not find an effect of high LSA-scores on task success. The mean LSA-score of the successful search queries was .26 and the mean LSA-score of the unsuccessful search queries was .24. LSA-scores do not predict the quality of search queries in our study very well. We will illustrate this with the task about Columbus (see Table 5.11). The query 'Columbus' had the highest LSA-score, but did not lead to success, because the answer to the search task was not mentioned on the content page about Christophorus Columbus, but on the content pages about 'Discoveries' or about 'Vikings'.

Table 5.11: LSA scores of the most frequently entered search queries for the search task about Columbus.

Search task	Frequent search queries (N)	LSA-score	Successful?
"It is often said that Christopher Columbus discovered America. Some say that this is not true. They say that another nation discovered America 500 years before Columbus did.Can you find out what people discovered America before Columbus did at Junior Winkler Prins online?"	- America (19)	0.27	No
	- Discovery America (17)	0.27	Yes
	- Columbus (10)	0.30	No
	- Christopher Columbus (7)	0.33	No
	- Vikings (3)	0.13	Yes

Quality of search results' selection

When we look at the behaviour of the children in selecting search results, we first of all see that most children tended to immediately choose the first search result from the list. From the search attempts in which children actually selected a search result (N=216), 110 selected search results were the first result from the results list. Further, almost all selected search results (211) were selected from the first 10 results that were presented in the results list. This tendency to select the first search result or a search result from the first 10 search results presented was also reported in previous studies (Bilal, 2000; Druin et al., 2009).

The children were not inclined to select more than one search result from the search results that were presented for one search query. Most children only selected one search result (86% of the search attempts). Two search results were selected within 11% of the search attempts and more than two search results were selected within 3% of the search attempts.

Of all analyzed search tasks that were conducted with the search engine (N=190), 83% (after one or more unsuccessful search attempts) were successful and led to the

right content page. With 96% of the selected search results no help was provided by the test instructor. From our analysis it is not entirely clear what the main reason is for the unsuccessful search tasks. A possible explanation is the fact that often children chose high-scent incorrect links. However, our research was not concerned with the quality of the search engine and therefore we did not further analyze the relevancy of the search results provided by the search engine.

Navigation and browsing categories

In this section, we will analyze children's navigation behaviour. What makes browsing successful or unsuccessful? Do children experience the same problems with browsing as reported in previous research or not? For this qualitative study of selected main and subcategories and content pages, the navigation paths for the five search tasks of 46 children were analyzed that were either successful or unsuccessful in completing the search tasks by browsing the categories[12]. We will describe problems with the layout of the websites, problems with the information structure of the websites and operational navigation problems. The problems presented are not an exhaustive list of all problems, but give an illustration of the most important navigation problems.

Problems with the location of the search box

The most important problem in our view with the layout of the websites, and particularly with the Image map website, was the location of the search box. Many children did not use the search engine when it was provided. We do not know for sure whether these children were not familiar with using a search engine or whether they did not notice the search box at all. The unconventional layout of the Image map website and the unconventional location of the search box in the top left corner may have caused the search engine to have been unnoticed by many children (see Figure 5.10). We reported in the corpus study (Chapter 3) that 43% of the websites with a search engine, presented the search engine in the top right corner of the page and 24% presented the search box in the top center of the page. Therefore, the top left corner is a non-conventional location for children's informational websites. However, on the other two versions, the search box was also placed at the top left side, but on these two versions more children used the search engine when offered than on the Image map version.

Problems with the layout

Most problems with the layout were experienced with the Image map website, most likely because of the unconventional interface design. When parsing the home page,

12. Amy Mooij (MA student) assisted by conducting qualitative analyses of the children's navigation paths.

children directly focused on a particular part (mostly on the main menu) of the Classic website and the Classical play website. This was in contrast to children's parsing behaviour on the Image map. Most children parsed the entire home page of the Image map and all main category links were looked at extensively. This was caused by the fact that the main menu covers the entire home page of the Image map. There were many problems with the main category links on the Image map website, because children had to 'mine sweep' the images to see the verbal link labels. However, many children did not directly recognize the main category links as such and did not notice their click ability, as can be seen in Figure 5.10.

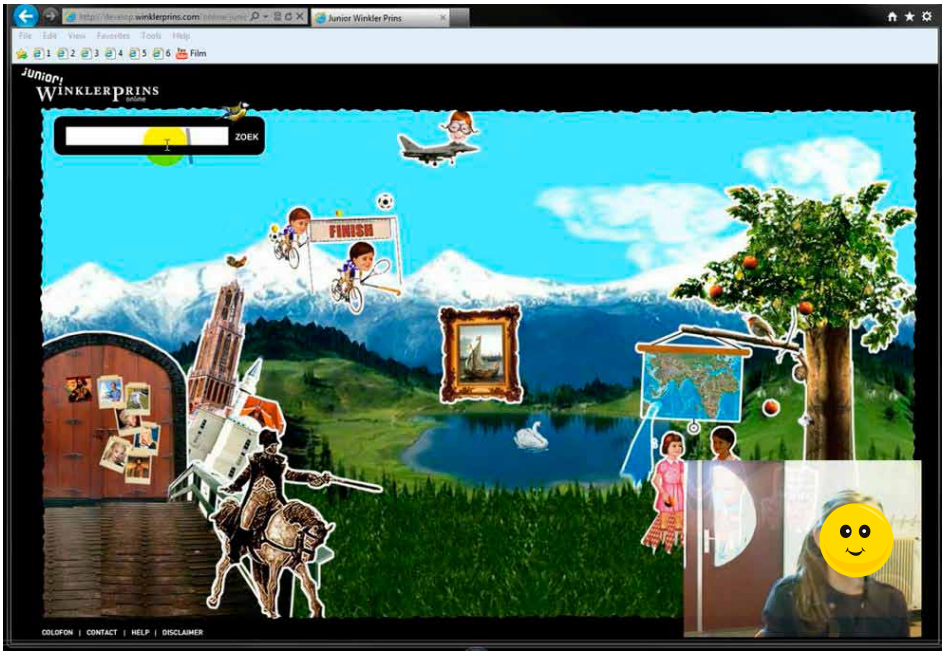


Figure 5.10: The unconventional location of the search box and the main category links as images at the Image map website

Problems with the information structure

Problems with selecting the right main and sub categories occurred on all three type of websites, but mostly on the Classic and the Image map websites. Choosing the right main category was especially problematic for many children, as was also reported by Hutchinson *et al.* (2006). Often, children had wrong expectations of the content behind main category links. For example, the main category link label 'human' was often selected for tasks about the first astronaut or about Columbus, because humans were mentioned in the task description. However, in fact the main category 'human' was about the human body.

Children did benefit from the addition of images with the main and subcategory link labels at the Classical play website in comparison to the Classic website (see Figures 5.11 and 5.12). The images helped children to choose the right main and sub categories and the children used more trial-and-error in choosing the right categories on the Classic website.

The addition of images on the Image map website often had an opposite effect on children's navigation behaviour, because the images without the verbal link labels were often not obvious (see Figure 5.10). In other words, addition of images to the category labels helped children in navigation, but corresponding text labels with the images were essential for quick understanding of the link labels.

Operational navigation problems

Children needed more assistance in operating the navigation tools on the Image map than on the other website versions. The sub category link labels on the Image map appeared in a pop-up layer across the home page and often children did not know how to get out of this pop-up layer (see Figure 5.13). There were no navigation tools provided to go back to the home page, as was the case at the other two website versions. Children had to find out for themselves that they had to click somewhere on the screen next to the pop-up layer to make the pop-up layer disappear.

Children's navigation skills

Many problems that children experienced on the websites cannot be attributed to the design characteristics of the websites, but are caused by children's own navigation and information skills. Differences between children's navigation and information skills are more a determinant for differences in search performance than differences between Web design characteristics. We will illustrate this by describing differences between search behaviour of children that were successful and unsuccessful in finding the right information on the websites.

Where successful children showed a clear pattern in link selection, less successful children lacked such a clear pattern. Less successful children especially had problems with the interpretation of the main category link labels. They often selected high-scent incorrect links and they almost never selected low-scent correct links. They often lacked domain knowledge to select a relevant category link or did not keep the initial search task in mind. Also, less successful children often did not recognize when they had selected the wrong main category link and got lost in the website's menu structure.

Successful children did keep the initial search task in mind and had fewer problems with interpretation of the main category link labels. They were well aware when they had selected a wrong main category. They were often able to select low-scent correct links and followed the optimum navigation path more often than unsuccessful children.

Figure 5.11: The main and subcategory labels without addition of images on the Classic website.

Figure 5.12: The main and subcategory labels with addition of images on the Classical play website.

Figure 5.13: The sub category link labels on the Image map website appear in a pop-up layer across the home page.

Browsing strategies

Successful children used both a trial-and-error strategy and a 'think-then-act' strategy on the Classic website. On the Classical play website, successful children only used one strategy: the 'think-then-act' strategy. This might be explained by the fact that the link labels on the Classical play website were presented with a picture that visually represented the textual link label, allowing children to make a better interpretation of the meaning of the link labels. This made the 'trial-and-error' strategy less necessary.

Less successful children exclusively used a trial-and-error strategy to find information on all three website versions. They often only used one strategy and had no plan B, when plan A did not lead them to the right content page. They had trouble keeping the search task in mind and often applied a 'loopy' navigation style, which means that they selected the same incorrect main and sub category links again and again. This looping navigation behaviour was also reported by Bilal (2000).

Another problem that occurred with some children while using the Image map website was that they were hesitant to click on the main and sub category links. The playful layout of the Image map website made them insecure to click on the links.

Processing the content page

Successful children often asked for assistance in understanding the content. Less successful children did not ask for help, but just gave a wrong answer. Reasons for giving

wrong answers were because of a wrong interpretation of the text, not recognizing internal hyperlinks to a relevant content page, because of low literacy skills, scanning the text too quickly or by not keeping the initial search task in mind. Successful children mostly had the capacity to reflect on the initial search task and made sure that they kept the search task in mind.

Processing problems

It was noted that children experienced more problems on all areas of navigation on the Image map than on the other website versions. Children even had trouble understanding the search tasks correctly when using the Image map website, which did not occur with the other website versions. This might be explained by the 'cognitive load theory' (Sweller, Merrienboer & Paas, 1998). Processing the Image map website might have taken so much cognitive energy from the children, that there was too little cognitive energy left to interpret the search task.

Game-experience

Finally, we found some clues that game-experience influences children's navigation behaviour. Children with little game-experience (mostly girls according to the results of the profile survey) often think first before they act. Children with a lot of game-experience (mostly boys) use the trial-and-error search strategy. This trail-and-error strategy is provoked when children do not know exactly where to go, which is often the case on the Image map website. Therefore, the use of the Image map seems more suited for children with a lot of game-experience.

5.6 Conclusions and discussion

RQ - chapter 5-1: Effects of interface design *(related to RQ 2)*

In this section we will answer our research questions.

1a. The first research question is about the effects of differences in interface design. We did not find any effects of the different design types on children's search performance on informational websites. The variation in design of the websites did not have an effect on search success and efficiency.

- Hypothesis 1a is rejected: there was no significant difference between children's search performance on the Classic, Classical play and Image map websites. The discrepancy between these results of the experiment with the results of the explorative study is most likely caused by the fact that in the explorative study, we did not take into account the variance between the children. This variance turned out to be a stronger determinant for the differences between

children's performance on the different website types than the difference in interface design of these website types (as discussed in Section 5.5.2). Also, the differences between the types of web design were not controlled for most factors in the explorative study, as was the case in the experiment. We will discuss this further in Chapter 6 of this dissertation.

1b. As opposed to the performance scores, there were differences in the subjective scores measured. Children were most positive about the Classical play website according to their scores on the SAM-scales for valence. This positive score is most likely not based on pragmatic issues, because both children that were successful and children that were unsuccessful in finding the right information on the Classical play website, gave this website higher scores than they gave the other two website versions. Apparently, their feelings about the website was based on more hedonic issues, such as whether they were attracted to the interface design of the website. Surprisingly, we saw the same pattern for children's evaluations of goodness: both successful and unsuccessful children evaluated the Image map version as less good than the other two website versions. There were no significant differences between children that were successful and children that were unsuccessful in finding the right information. Apparently, search success was not a determinant for children's evaluation of goodness.

   ► Hypothesis 1b is accepted for the most part: playful 'visual' design (i.e. the Classical play website) had a positive influence on children's evaluation and playful 'navigation' design (i.e. the Image map website) had a negative influence on children's evaluation. However, this evaluation was not related to children's search success, because both successful and unsuccessful children gave the same evaluation scores for the various websites.

RQ – chapter 5-2: Relation between performance and evaluation *(related to RQ 3)*

We did not find evidence for a strong relation between affectivity and usability in our study. Children's affective responses were not based on the effectiveness of search performance on the websites, as we hypothesized based on the results of the explorative study. However, their affective responses could have been based on pragmatic issues besides the final success in finding the information. For instance, their affective responses could have been based on the ease of use while interacting with the interface, as was reported by Sim *et al.* (2006) who found that children appeared to have less fun when their interactions had more usability problems. However, in our study children's affective responses towards the search systems seem to have been independent from their actual search behaviour and most likely based on perceived hedonic quality and aesthetics of the interface. This could be best tested by a pre and post measure of affective responses, to see whether actual behaviour changes children's attitude to-

wards product evaluations such as beauty, goodness and fun (Van Schaik & Ling, 2008). Because of time constraints it was not feasible in our experiment to conduct both pre and post tests to measure affective responses.

We did find some proof that the children's affective responses were based on the ease of use while interacting with the interface (i.e. pragmatic quality). An analysis of verbal utterances of 25 children that used the Image map website showed that most of the utterances were negative and related to the websites' low pragmatic quality. This result supports the fact that the children's valence scores and evaluation scores for goodness were lowest for the Image map website and that these scores were based on perception of pragmatic quality. The fact that almost no hedonic utterances were made, supports the fact that playfulness and expressive aesthetics did not have a large influence on the children's attitude towards and evaluation of the website. These qualitative results correspond with the results of the explorative study in which we also found that positive, hedonic, emotional expressions were less important for children when evaluating the search interface than perception of pragmatic quality (i.e. usability) (see Section 4.7).

The fact that we did not find a relation between beauty and usability in our experiment can also be caused by children's tendency to indicate the highest score on the scale (Markopoulos et al., 2008). Most children felt very positive about the three websites and little frustration was uttered. Therefore, we did not find evidence that pragmatic frustrations lower children's perception of aesthetics as Tuch *et al.* (2012) found in their research with adults. This was most likely also caused by children's tendency to give socially desirable answers. Although they were often not successful at all in conducting a search task, children found it difficult to be negative about the system and tended to blame themselves instead of the system. This was also reported by Serenko (2007), who studied the self-serving biases of interface agent users. He found that adult users may attribute their success to an interface agent and hold themselves responsible for task failure, just as the children may have done in our experiment.

- Hypothesis 2 could not be accepted: in our study children's affective responses towards the search systems seemed to be independent of their actual search behaviour. High performance scores (i.e. search success) did not have influence on children's evaluation scores in our experiment. Evaluation scores were most likely based on perceived hedonic quality and aesthetics of the interface. However, a qualitative analysis of children's emotional expressions uttered while using the Image map website did provide proof for hypothesis 2 that a positive perception of pragmatic quality (regardless of actual success in conducting the search tasks) has a stronger influence on children's overall evaluation of a search interface.

RQ – chapter 5-3: Interaction with the chosen search strategy *(related to RQ 4)*
From this experiment we can conclude that the search strategy that was used by the children was much more a determinant for their search performance than the interface design of the search environment. Searching with a search engine proved to be much more effective and efficient than browsing the navigation structure. This was the case for all three types of website design in our experiment.

We found that there was a significant difference in success scores when children received help from the test instructor. However, this difference meant that children were less successful. In other words, children did not become successful because of the help they received. These findings prove that help was only offered to motivate and reassure children during the search process (as recommended by Markopoulos et al., 2008) and that the help offered did not have a significant effect on children's success in finding the right information.

It turned out that searching was more effective and efficient than browsing the categories. This is quite logical when considering the fact that children nowadays are members of the 'Google-generation' (Rowlands, 2008). However, we did not look at the long-term effects of the fact that the children prefered and were better in searching than in browsing. Searching instead of browsing might, for example, have a negative effect on children's knowledge of information taxonomies. The search engine can be compared to a 'black box' that does not give insight into how information is related in a taxonomy. Future research is needed to study long-term effects on children's knowledge of information architectures of searching with a search engine in comparison to browsing categories.

Search and navigation behaviour on the Web is constantly changing. Currently, traditional search engines as Google are losing 'traffic', because people are using social media such as Facebook and Twitter as their primary Web entrance to find information more and more (Xiang and Gretzel, 2010). This might also cause changes in children's search and navigation behaviour.

- Hypothesis 3 is partly rejected and partly accepted. There was no difference in the children's search performance by browsing the main categories between different web design types (this part of the hypothesis is therefore rejected). There was also no difference between children's search performance on the different interfaces by using the search engine (this part of the hypothesis is therefore accepted). Overall, the children performed better by searching with a search engine than by browsing the main categories (this part of the hypothesis is therefore accepted).

RQ – chapter 5-4: Problems and successes with keyword searching and browsing *(related to RQ 4)*

When studying children's search performance in more detail, we found that many problems that children experienced on the websites could not be attributed to the design characteristics of the websites, but were caused by the children's own lack of navigation and information skills. In other words, children's navigation and information skills are a better predictor of children's search success than the type of design characteristics of the search interface. Children with a lot of internet experience encountered fewer problems with searching and browsing than children with little internet experience. Also, children with a lot of domain knowledge were more successful in formulating relevant search queries, selecting relevant search results, or selecting relevant main and subcategories. Children that were able to check and monitor their own activities, termed metacognition (i.e. the voluntary control an individual has over his own cognitive processes) (Brown & DeLoache, 1978) were more capable of keeping the initial search task in mind, or of recognizing when they selected the wrong main or sub category link.

An important lesson learned from this experiment is that the variance between children is much more a determinant for differences in search success and performance than the variance between search interfaces. This makes the challenge for designers to design interfaces that support children in effective information searching even greater. It also stresses the importance of educating children in navigation and information skills. However, the most important lesson learned from this experiment is the fact that children's search success and search efficiency was much larger when they used a search engine, than when they browsed the main categories of the website.

RQ – chapter 5-5: Exploring research methods

The final research question in this chapter concerned the suitability of existing methods for research with children to measure feelings and perceptions of pragmatic and hedonic quality that are used in research with adults. We experienced that methods such as the SAM-scale for emotional arousal and the AttracDiff 2 questionnaire are not valid methods in research with children. Unlike Greenbaum *et al.* (1990), we could not validate the use of the SAM-scale for emotional arousal with children. These opposite results might be caused by the different settings in which the method was used. The pictures in the SAM-scale for emotional arousal might be related easier to 'fear of the dentist' by children (which was the case in the research of Greenbaum et al., 1990) than the feeling experienced with using a website. We could, however, validate the SAM-scale for emotional valence with children.

Almost all picture anchors for the semantic differential scales for the pragmatic and hedonic items could not be validated for use with children in our experiment (which we did not expect based on our pilot test of the semantic differential scales with 14

children). The children in the experiment did not interpret the picture anchors as intended by the developers of the AttracDiff 2 questionnaire (Hassenzahl et al., 2003). The children only made a clear distinction between which end of the scale was intended as positive and which end of the scale as negative in their opinion. The instrument could not make clear to the children what the difference is between perceptions of hedonic and pragmatic quality of the websites. The pragmatic and hedonic items were all seen as the same construct of valence (i.e. positive versus negative).

Although we did find differences in scores for the product evaluations beauty, goodness and fun, we doubt whether children were able to make a clear distinction between these concepts, as was also reported by Read *et al.* (2002). Fortunately, we were able to use an objective method to measure emotional arousal by using the Q-Sensor. This method turned out to be valid to measure emotional arousal with children in contrast to the subjective methods used in our experiment.

Based on this experiment, we cannot judge whether the UX models presented in research with adults (Hassenzahl, 2004; Thüring & Mahlke, 2007; Van Schaik & Ling, 2008) can also be applied to children's UX. Although we could not entirely reproduce UX research concerning perceptions of pragmatic and hedonic quality in our experiment, we did show the difficulty of reproducing UX research with children. More research is needed to develop valid methods to measure children's perceptions of pragmatic and hedonic quality.

ZOEK

Chapter 6

Conclusions and discussion

6.1 Introduction

In the introduction to this dissertation, we presented our 'triple jump' research design and in the chapters following, we reported the results of the three iterative research steps taken. We can conclude that we have reached our research goals by taking the three research steps. Every step led to other steps, which proves the value of our iterative research design. Figure 6.1 shows an overview of our research steps, with a short summary of our main research findings for each step.

We will start this final chapter by summarizing our main research findings per research question. After that, we will reflect on our research findings, first, from the perspective of the model presented at the end of Chapter 2, and after that, from the perspective of the research methods used in this research project. In Section 6.4, we will give recommendations for web designers and educators, because some of the lessons learned from our research can be applied by those who supply children with interfaces or by those who educate children in using these interfaces. In the final section of this chapter, we will discuss the implications of our research findings for trends in current visual design and search engine design and we will discuss future research.

Step 1 (*hop*):	Step 2 (*step*):	Step 3 (*jump*):
Corpus study We identified design conventions for informational websites for children and we identified three types of children's websites based on a cluster analysis of the analyzed design characteristics: 1: Classic 2: Classical play 3: Image map	**Explorative study** The results of the explorative study suggest that children's search performance is most effective and efficient with the classical play website type and least with the image map website type. Although the number of expressed positive hedonic emotions was highest with the Image map website, children's overall appreciation was lowest for the Image map website. The results also suggest that perceived pragmatic quality is more important for children than hedonic quality.	**Experiment** In our experiment, we could not accept all hypotheses based on the results from the explorative study. We did not find any effects of the differences in design on objective children's search performance scores. However, we did find effects on subjective affectivity scores. Children gave most positive scores to the Classical play version of the website and evaluated the Image map version as least good of the three versions. Their feelings and evaluations were not based on final success of finding the information, but they were more related to usability than to beauty and fun. Most important result is that searching proved to be much more effective and efficient than browsing the categories.

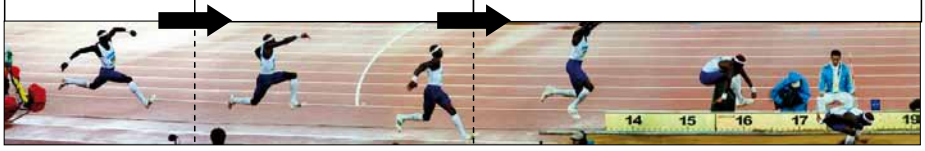


Figure 6.1: Main research findings for each research 'step'

6.2 Summary of results

We will now answer the main research questions in more detail. Where relevant we will give an overview of the results of the explorative study and experiment and discuss the agreements and discrepancies between the two studies.

RQ 1. Which design conventions on interfaces for children are specific for children and which of them are general design conventions?

Most design conventions that were identified in the corpus study of children's informational websites reported in Chapter 3 correspond to general web design guidelines. Most of these identified conventions also correspond to the literature about how children interact with websites (Nielsen & Gilutz, 2002; Meloncon et al., 2010). Design conventions such as the availability of a main menu as the primary navigation tool, use of short texts, high contrast between text and background, marking of hyperlink click-ability, conventional location of items such as the logo (top left of page) and service links (bottom of page), ordered page layout, a solid background and a consistent layout through the website do not only prove to be helpful for adults but also for children (Nielsen & Gilutz, 2002). Apparently, designers of children's informational websites follow general web design guidelines when designing interfaces especially for children.

Still, many websites in our corpus showed design characteristics that do not belong to a set of generally accepted design conventions. Apparently, designers often do not agree about which type of design characteristics is suited for children's search behaviour. For example, 58% of the websites required scrolling up and down the page to view all of the information available. Meloncon *et al.* (2010) do not provide conclusive proof in their study about children's preferences regarding page length and their willingness to scroll, while Nielsen and Gilutz (2002) recommend no scrolling on web pages for children, based on their research results.

A closer look at the data of the corpus revealed three categories of children's informational websites with specific design approaches for children: a Classic, a Classical play, and an Image map design approach. The layout of the Classic websites (see Figure 6.2) is kept minimal; key elements of the pages are the center of attention and page components are located at conventional locations as recommended in the literature on general usability guidelines for web design (Koyani et al., 2006; Nielsen and Tahir, 2002). Further, the Classic websites are characterized by a solid background color (mostly white), little clutter on the pages and low screen density. The idea behind this is that on Classic websites, children will not get lost easily because of the consistent menu structure and layout throughout these websites. They also contain low number of images, pictures, illustrations and animations.

The Classical play websites (see Figure 6.3) are characterized by a traditional arrangement of information, just like the Classic websites, but the visual design of this type of

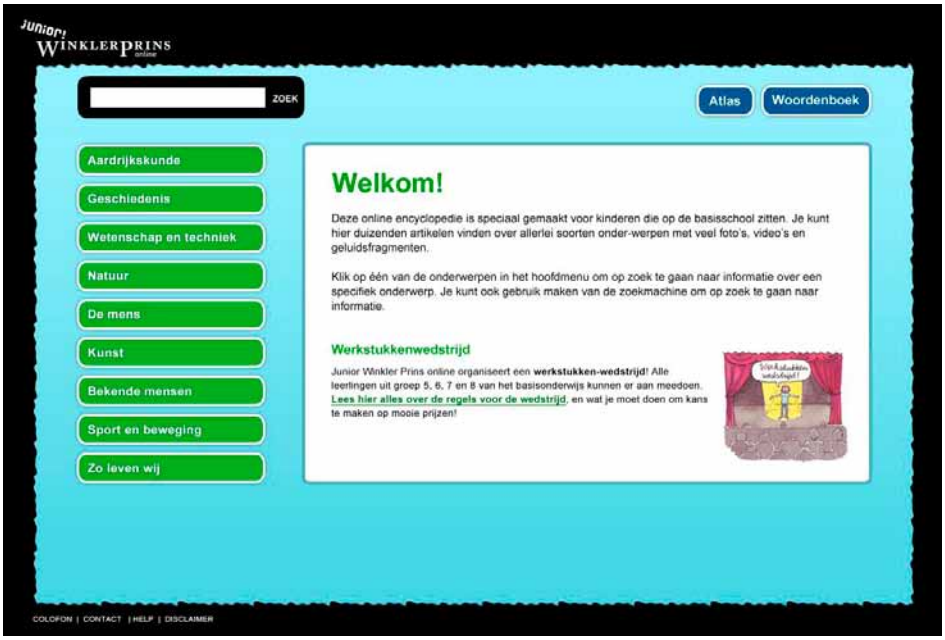


Figure 6.2: Classic website type

Figure 6.3: Classical play website type

website is playful, just like the Image map websites. The Classical play websites present a substantial number of images, animations and colors. Menu labels in the websites of this type are accompanied by icons. Menu structures and layout on the Classical play websites are consistent throughout the pages and many items are presented at conventional locations. Another common characteristic of this group of websites is the presentation of a clear focal point on the web page, often in the form of an image or animation, which immediately draws the user's attention.

No classic web design characteristics are used on the Image map websites (see Figure 6.4). The visual design and navigation of the websites in this group are based on image maps that incorporate objects or locations that children know from real life or from fiction on which a child can explore in search of specific information. This makes information search on these websites a playful experience. The use of multi-media is high on these Image map websites and standard components of a website, such as the main menu or a search box are incorporated into the visual metaphor of the Image map at a non-traditional location. The background of these websites is mostly non-solid, because of the visual metaphor that covers the entire screen. Search engines are less common in the group of Image map websites.

The following steps in our research project gave us more insight into which design characteristics are suited for children's search behaviour and which characteristics are not suited. We found that there are several design conventions that are relevant for all children's informational websites as listed at the beginning of this paragraph. Furthermore, based on other design characteristics - that were not conventional for all analyzed websites - we identified the three categories of children's informational websites. These

Figure 6.4: Image map website type

three categories of children's informational websites might have different effects on children's search performance and perceived pragmatic and hedonic quality of the interfaces, which was an important issue for the following steps in our research project.

RQ 2. What is the effect of design characteristics of informational websites for children on their search performance and on their appreciation of the search interface?

In the explorative study reported in Chapter 4, children's search performance scores were lowest on the Image map website and highest on the Classical play website. Therefore, we hypothesized that playful 'visual' design has a positive influence on children's search performance and playful 'navigation' design has a negative influence on performance. We tested this hypothesis in the experiment reported in Chapter 5. However, we did not find evidence for this hypothesis. There was no significant difference between children's search performance on the three different types of interface design. But how can we explain this discrepancy between the results of both studies? Differences between the research results may not only have been caused by the differences in the research stimuli used, but also by the research methods used or by the variance between the children that participated in the studies.

Concerning the research stimuli used, the differences found in the explorative study between the different types of website design could be caused by other factors than only visual and navigation design, such as formulation of main and subcategories or the working of the search engine. In the experiment, we worked with controlled stimuli and differences caused by the research stimuli could only be caused by the manipulation of visual and navigation design.

Furthermore, it turned out in the experiment that the variance between children is much more a determinant for differences in search success and performance than the variance between search interfaces. We did not take into account between-children variance and residual variance in the explorative study, as we did in the experiment. In conclusion, the results in the explorative study could very likely be caused by uncontrolled differences between the interfaces and by variance between the children.

We will now discuss the results of the explorative study and the experiment in more detail. We first explored research question 2 by studying children's search performance on three existing, Dutch, informational websites for children (Classic: schoolbieb.nl / Classical play: willemwever.nl / Image map: Kids.kennisnet.nl). The results of the study showed that search performance was most effective and efficient with the Classical play website type (see Figure 6.6). Close behind were the children's performance scores on the Classic website (see Figure 6.5), which showed that search performance was also quite effective and efficient with the Classic website type. The children also evaluated the Classical play website as easiest to work with. The results for the Image map website (see Figure 6.7) were negative: search performance was least effective and efficient while using the Image map website and the children evaluated the Image map website

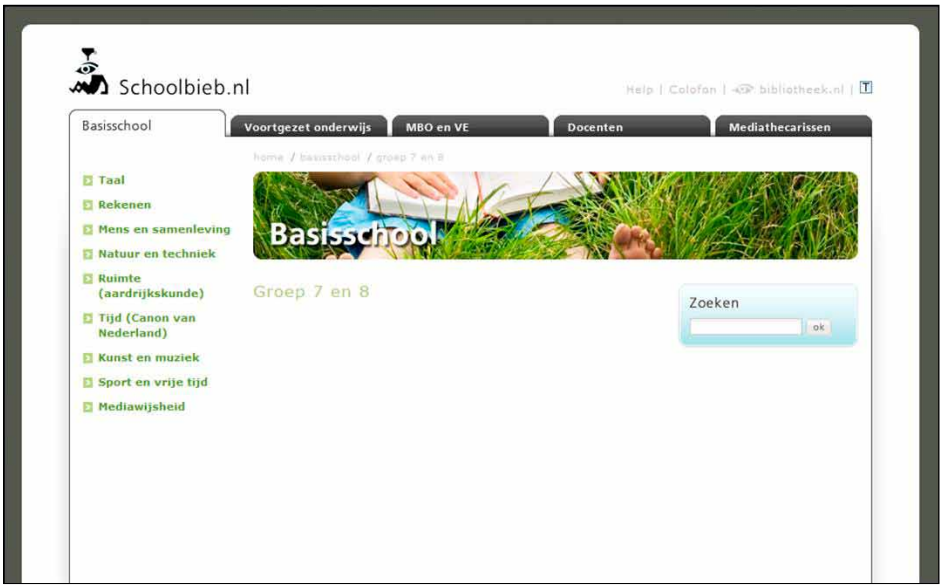


Figure 6.5: Classic website (from Chapter 4)

as most difficult to work with. These results showed that the Image map website type was least suited to support children's search performance.

Besides the actual performance scores on the websites, we also explored the user-centered experience with the websites by analyzing emotional utterances expressed while conducting the search tasks. We found that most emotional utterances expressed in the study concerned the Image map website. Most of these emotional expressions were based on pragmatic attributes, but also many expressions were based on hedonic attributes. Apparently, the playful design approach of the Image map website evoked a lot of emotional expressions, both pragmatic and hedonic.

The Classic website evoked mostly negative emotions, solely based on pragmatic reasons. Also the few positive emotions expressed while using the Classic website were based on instrumental qualities, such as satisfaction or relief in finding the answer to the search task. Concerning the Classical play website, most emotional expressions were also based on pragmatic reasons. There were some emotional expressions based on non-instrumental qualities, such as fun, tingling or astonishment about the playful elements, but they were rare.

For the overall evaluation of the websites, pragmatic quality turned out to be more important for the children than hedonic quality. However, relatively more emotional expressions concerning the Image map website were based on hedonic reasons than concerning the other two websites. It seems that aesthetic design of the Image map website and the fun experienced by the playful elements on this website stimulate children in expressing emotions. However, these emotions did not appear to have con-

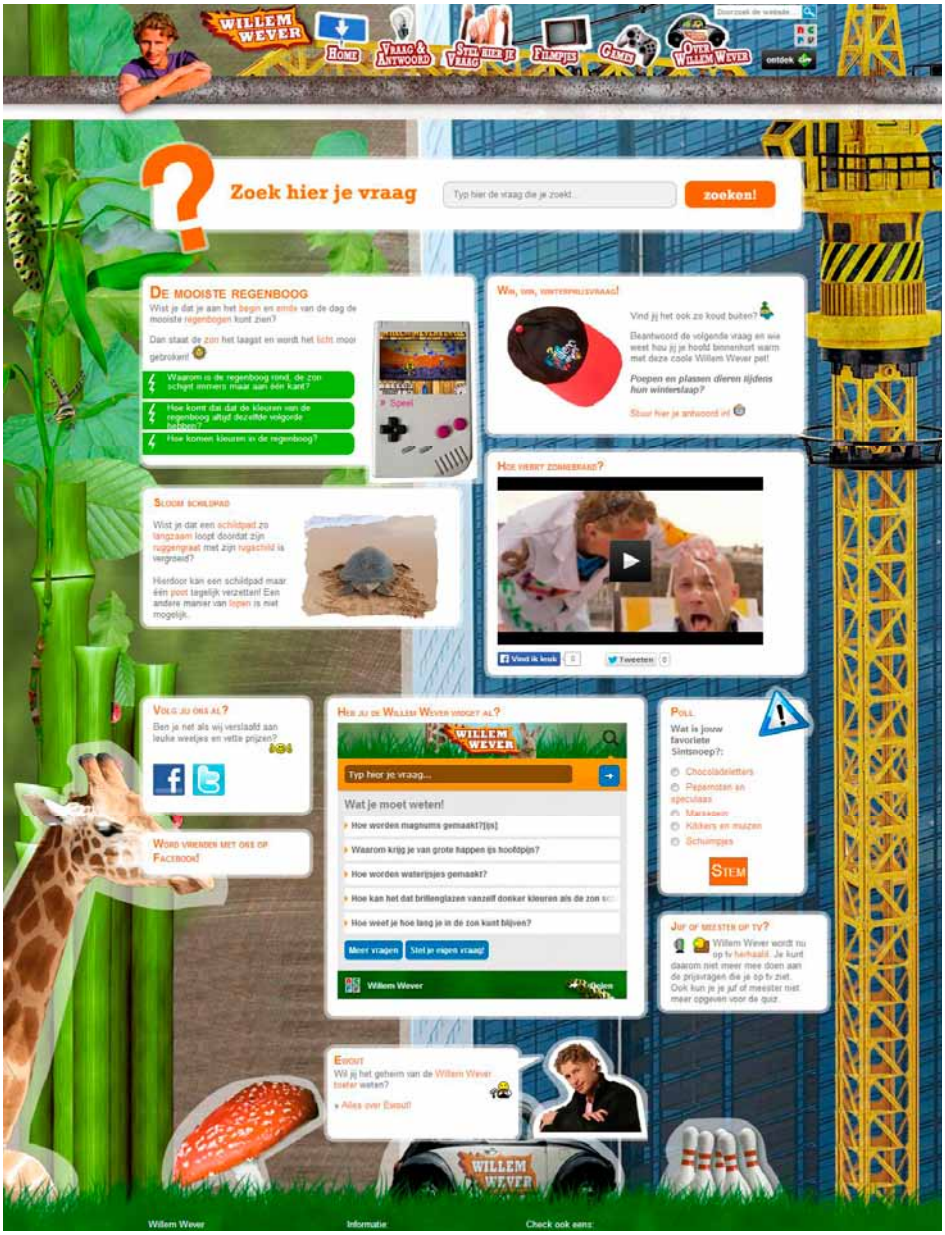


Figure 6.6: Classical play website (from Chapter 4)

sequences for their overall evaluation of the website. In contrast, no emotional expression about the Classic website was based on hedonic issues, but overall evaluation was highest for the Classic website. It seems that hedonic issues do provoke emotional expressions, but these are not of importance for children when evaluating the interface for future information search tasks.

Figure 6.7: Image map website (from Chapter 4)

In a follow-up study, we empirically tested whether the results of the explorative study could be validated. It could also be the case that the findings were caused by the many uncontrolled differences between the websites in the explorative study, besides the differences that were specific for the three types of website design identified in the corpus study. Therefore, we set up an experiment in which the differences between the three website types were controlled by manipulating one website into three different versions with characteristics that were specific for the three types of website design.

Based on the results of the experiment, we rejected the hypothesis about the effects of the three different types of informational web design on children's search performance. The variation in design of the websites did not have an effect on search success and efficiency. As opposed to the performance scores, there were some differences between the three types of website design in the hedonic scores measured. Children were most positive about the Classical play website according to their scores on the SAM-scales for valence. This positive score is most likely not based on pragmatic issues, because both children that were successful and children that were unsuccessful in finding the right information on the Classical play website, gave this website higher scores than the other two website versions. Apparently, their feelings about the website are based on more hedonic issues, such as whether they are attracted to the interface design of the website. However, this result is difficult to reconcile with the fact that there were no significant differences between the children's perceptions of hedonic quality (i.e. unruly vs. manageable) and between their evaluations of fun and beauty of the three websites. These results do not support any influence of hedonic issues. Interestingly, we saw the same pattern for children's evaluations of goodness: both successful and unsuccessful children evaluated the Image map version as less good than the other two website versions. There were no significant differences between children that were successful

and children that were unsuccessful in finding the right information. Apparently, search success was not a determinant for children's evaluation of 'goodness'.

The results of the explorative study about the effects of playful design characteristics on children's search performance were falsified in the experiment. We did not find differences in search performance caused by the differences in design characteristics in the experiment. The earlier results were probably based on other differences between the three websites than the playful characteristics, for example, differences in menu structure and formulation of headings. In the explorative study, differences in menu structure and formulation of headings were not controlled for, while we did control for these differences in the experiment.

However, the results of the final evaluation of the three website types were similar in the two studies. The children in the explorative study were negative about the Image map website (by ranking it unanimously as the last one to use in future search tasks), just like the children in the experimental study. The scores in the experiment for emotional valence and evaluation scores for goodness were lowest with the Image map website in comparison to the Classic and Classical play website. Hence, the results from the explorative study and the experiment clearly converge in this respect that the Image map website was evaluated as least good and least attractive for conducting future search tasks.

A remaining question is whether these results were based on children's perception of pragmatic or hedonic quality. In other words, were children less positive about the Image map website because of the usability problems they experienced, or were children more positive about the other websites because of the aesthetics or fun experienced?

The results of the explorative study clearly showed that hedonic issues do provoke emotional expressions, but that these are not of importance for children when evaluating the system for future information search tasks. The results of the experimental study were less clear about this, because final search success did not appear to influence children's emotional valence or evaluation of goodness of the website. In the experiment both successful and unsuccessful children gave the same evaluation scores for the different website types. Therefore, unlike the results of the explorative study, evaluation scores in the experiment could be based on perceived hedonic quality and aesthetics of the interfaces instead of on pragmatic quality.

However, a qualitative analysis of children's emotional expressions uttered by 25 children while using the Image map website in the experiment did provide proof for the hypothesis that a positive perception of pragmatic quality has a stronger influence on children's overall evaluation of a search interface. As discussed in Section 5.4.4, this smaller group of children was representative for the total group of children that participated in the experiment.

Regardless of actual success in conducting the search tasks, children's evaluation scores in the experiment were more positive about the Classical play website and less positive about the Image map website. And there were no differences in evaluation scores for beauty and fun in the experiment. Therefore, we strongly believe that children's evaluation scores are influenced by instrumental qualities and perception of pragmatic qualities (i.e. usability) during task performance. In other words, when children can easily work with an interface and the interface is user-friendly, their evaluation scores are more positive about the interface than when the interface causes problems during navigation and information search, regardless of the final success or failure of finding relevant information.

We conclude that playful navigation design did not have a positive influence on the children's search performance or final evaluation of a search interface. Playful visual design did have a positive influence on the children's perception of hedonic quality, based on their positive emotional expressions about the hedonic quality of the playful elements. However, this is of less importance for children when evaluating a search interface on goodness or for future use. The Image map website scored lowest for children on emotional valence, goodness and for future use. The Classical play website scored highest, closely followed by the Classic website. Probably because of its high pragmatic quality and usability.

RQ 3. What is the relation between children's search performance and their affective responses towards the search interface?

To our knowledge, research on child computer interaction does not discuss the relation between performance and affective responses as was reported with adults in research by Tracktinsky *et al.* (2000), Hassenzahl (2004) and Van Schaik and Ling (2008). Therefore, we will first discuss our research results concerning child users in light of previous research concerning adult users. Tractinsky *et al.* (2000) concluded that a product's beauty is a stronger indicator for its perceived usability than its actual usability, because post-use usability ratings were not affected by actual usability. Moreover, they claimed that "what is beautiful is usable". Beauty of an interface, in their view, overrules all other interface characteristics and therefore influences users' overall evaluation of the system. In our research, we falsified this statement. For the children in our studies, beauty was no indicator for perceived usability at all. Unfortunately, we did not measure pre and post-use usability ratings. However, goodness and overall evaluation of the search systems were more influenced by perceived usability than perceived beauty.

Our results are more in line with the results reported by Hassenzahl (2004). He found that goodness was more closely related to attributes of pragmatic quality (e.g. a product is perceived as simple, practical, clear, predictable, etc.), as was also the case in our research. Van Schaik and Ling (2008) also conducted an experiment on the interplay between components of User Experience. They found that all measures (i.e. evaluation

of goodness, attributes of hedonic and pragmatic quality, task performance and mental effort) except evaluation of beauty, were sensitive to manipulation of web design. Evaluation of beauty was influenced by hedonic attributes (identification and stimulation), but evaluation of goodness was influenced by both hedonic and pragmatic attributes as well as task performance and mental effort. In retrospect, we doubt whether children could make such a clear distinction between evaluation of beauty and goodness in our study, because these constructs are quite similar in the perception of children, as was also reported by Read *et al.* (2002). In their research with the Fun-Sorter in which children needed to sort products by concepts such as 'worked the best', 'liked the most', 'most fun' and 'easiest to use', they say: "This was quite difficult for the children with the result that some constructs turned out to be quite similar." Based on our research results, we support this claim that it is difficult for children to make a clear distinction between different constructs as beauty, goodness, and fun. Children's evaluation scores of these different constructs were also close to each other in our research.

Our research results do not corroborate the link that Hartmann, Sutcliffe, and De-Angeli (2008) found between aesthetics and usability. They found that when users' usability experience was poor, positively perceived aesthetics could positively influence overall appraisal of a system, suggesting "aesthetics could be an important determinant of user satisfaction and system acceptability, overcoming poor usability experience." This was certainly not the case with the Image map websites in our studies. Although children did express positive feelings about the fun and playful elements in the Image map website, these elements could not overcome children's poor usability experience with the playful navigation design. Furthermore, Hartmann *et al.* (2008) argued that the relative importance of aesthetics is related to the user's background and task. When the users' task is goal-oriented (i.e. when reaching the goal is the main focus) then usability factors will weigh more than aesthetic considerations. When the user's task is action-oriented (i.e. the interaction experience is more important than reaching the goal) users choose designs based on a general impression of aesthetics and engagement. However, based on our findings, we do agree with Hartmann *et al.* (2008) that when the user's task is goal-oriented, as was the case with the children in our studies, then usability factors weigh more than aesthetic considerations.

Finally, also Tuch, Roth, Hornbaek, Opwis and Bargas-Avila (2012) conducted an experiment on the aesthetics-usability relation and they also found - under certain conditions - evidence for the relation "what is usable is beautiful". They found that the frustration of poor usability lowers ratings on perceived aesthetics. In our research, it appeared to be difficult to measure perceived aesthetics with children. The frustration of poor usability did lower final evaluation ratings in our research, however, we cannot say whether it also lowered perceived aesthetics for the children in our research.

The research we have discussed so far did not concern children but adults. Sim, MacFarlane and Read (2006) did study children to validate the Fun Toolkit, a tool to evaluate technology with children. They tried to relate the constructs 'fun' and 'usability' in their study and report that children experienced less fun when there were more usability problems. They conclude that it is not all about fun for children and that usability does matter to them. Based on our research, we could not agree more with Sim *et al.* (2006): usability does indeed matter to children. It seems that when a task is goal-oriented - as was the case in our research on children's information interaction behaviour – there is not much difference between children and adults in that usability factors weigh more than aesthetic considerations (Sim et al., 2006; Hartmann et al., 2008).

RQ 4. What is the difference between browsing and searching in terms of efficiency and effectiveness of children's search task performance?

Based on our explorative study, it is not possible to deduce certainties about the difference between children's search performance by browsing categories and by searching with a search engine. On both the Classic (see Figure 6.5) and the Classical play websites (see Figure 6.6), almost all children used the search engine (29 of 32 children), which made a comparison between searching and browsing invalid. And almost all children (29 of 32 children) browsed the categories on the Image map website (see Figure 6.7), because the search engine was hard to find on the Image map website in the explorative study. Therefore, it was impossible to make a valid comparison between searching and browsing based on the results in this study.

In the follow-up experimental study, we controlled the search method by providing or not providing a search engine. Therefore, we were able to make a valid comparison between searching and browsing from our experimental data. We concluded that the search strategy that is used by children is much more a determinant for their search performance than the interface design of the search environment. Searching proved to be much more effective and efficient than browsing the navigation structure. This was the case for all three types of website design in the experiment and for all ages.

Not only are children more successful information-seekers when using a search engine than when browsing categories, the results of our explorative study also show that children prefer using a search engine over browsing the categories. This statement is true under the condition that the search engine gets noticed by the children. Many children that were offered a search engine on the Image map website in our experiment did not notice it. Apparently, the location of the search box is also critical for the use of the search engine. Furthermore, children's prior experience with using search engines to search for information, also determines whether children will use a search engine. These results correspond to the research results of Bilal (2000) in her research on children's use of the Yahooligans! Web Search Engine, but are in contrast to the research results of Druin (2003) in her research on the use of the International Children's Digital Library

(ICDL). By tracking the web logs, Druin (2003) found that, of 60,000 unique users between the ICDL's launch in November 2002 and September 2003, approximately 75% of the searches used category search (browsing), 15% used place search (by selecting a topographical location using a world interface) and just over 10% of the searches used keyword search. This difference might have been caused by the fact that on the ICDL there was much more visual emphasis on the main navigation tool to browse categories than on the search engine tool to use keyword search, which can also be seen on Figure 2.11. Our results are also in contrast to the research results of Hutchinson, Bederson and Druin (2006) who found that children are capable of using both keyword search and category browsing, but that they generally prefer and are more successful with category browsing. However, this research was also conducted with the ICDL with the same visual emphasis on the category browsing tool.

Further diagnosis of children's search behaviour (Section 5.5.4) supports the fact that searching is more successful than browsing categories. The children experienced far more problems by navigating and browsing categories than by using the search engine, such as problems with the layout of the website (in particular on the Image map website, because of the menu that was incorporated in an unconventional design), problems with the information structure, and operational navigation problems.

6.3 The research findings in light of the dimensions of the information problem solving activity model

At the end of Chapter 2, we gave a summary of the variables that would be studied in our research project, presented in the same manner as the dimensions of the IPS activity of Lazonder and Rouet (2008). We will now reflect on the results of our study in light of the model presented at the end of Chapter 2.

6.3.1 Contextual variables

We chose to focus the contextual variables in our research on an educational context, which means that we studied children in a school setting. We focused the search tasks on fact-based assignments about school-related subjects with the goal to collect relevant information, because we think that this type of search task is most representative for this context.

We have gained a good understanding of children's information search behaviour, that can be relevant for many situations. At first sight, our research findings are relevant for children's information search behaviour in school settings for school assignments. We cannot automatically generalize our research findings to other situations. Further research on the generalizability of our research results is needed to prove whether our

results are also relevant for children's digital information search behaviour in other contexts, such as at home or for children's self-imposed search tasks. We do think that our results are very valuable for designers of children's search interfaces to support children in effective information search in different contexts. In many other contexts, children's information-interaction will also be goal-oriented (e.g. "I want to find the game a friend told me about" or "I want to see Justin Bieber's latest music video"), which probably makes the principles of our research just as relevant for these other contexts in that usability factors weigh more for children in finding the information than aesthetic considerations.

6.3.2 Individual variables

An important lesson learned from our research is that the variance between children is much more a determinant for differences in search success and performance than the variance between search interfaces. Children's domain knowledge is most likely a better predictor of children's search success than the quality of the navigation structure or search engine. When children have more knowledge about a subject, they know which categories or search queries are relevant for the search task. We did not focus our research on the effects of children's domain knowledge on their search performance, which is an interesting topic for future research. Tanni (2008) reported that children's prior experience with search interfaces and their web skills are better predictors of search performance than the design of the interface. When a child has prior experience with searching with a search engine, he or she will notice or use the search engine sooner than a child that has no experience using a search engine. Our research results suggest that reading skills are better predictors of children's search performance than the formulation of the text labels and content on a web page, because children that had high reading skills automatically had less difficulties with reading and selecting text labels and reading and scanning content pages than children with lower reading skills. Further, metacognitive skills proved to be good predictors of children's search success, because children who monitored their search process (Lazonder & Rouet, 2008) and were focused on their goals and constantly related the steps in the process to their search goals, had less trouble finding relevant information than children with lower metacognitive skills. The focus of our research was not on effects of individual differences between children on their search performance. Therefore, based on our research results, we cannot make valid statements about the precise effects that these individual variables have on children's search performance. However, our research results point in the same direction as reported by Tanni (2008) that children's prior experience with search interfaces and their web skills are better predictors of search performance than the design of the interface.

6.3.3 Resources variables

Children's search performance in the explorative study was most effective and efficient with Google. The children also unanimously ranked Google as the most preferred search system to use for future information search tasks. A logical explanation for this result is that the children already had most experience with searching information via Google in comparison to the other websites used in the explorative study. We used Google as the benchmark for search engine interfaces in our research project, because Google has dominated as a worldwide search engine for years now. In 2010, Google's search market share was 91.4% (Data source: StatCounter Global Stats).

Children are successful using Google, because – unlike the average children's informational website - a Google-type search engine will always present search results, whatever query a child types in the search box. Also, Google helps with formulating search queries, because of the query suggestion tool that unfolds with suggestions of relevant search queries while one is typing a query. Furthermore, Google helps with the reformulation of a search query by providing spelling help when a query is not spelled correctly (i.e. "Did you mean…").

However, we also experienced the known problems with the fact that Google is an uncontrolled resource. Children had difficulties deciding whether search results were relevant for their search tasks and whether the information was provided by a reliable source (Druin et al., 2009). Furthermore, children were overwhelmed by the large amount of information that is provided by Google. In our experience, based on the results of the explorative study, Google does not always guide children in a safe manner to relevant and appropriate information for children.

The greatest effect found on children's search performance from the tested resources variables in our research, is the search method used by the children: that is browsing versus searching. When a child used the search engine, search performance was far more effective and efficient than when the child browsed the categories in search for relevant information. The children also experienced fewer problems using the search engine than with browsing the categories. Probably, the most important problem with the search engine was the location of the search box, because we experienced that when it was not placed at a conventional location (as was the case in our experiment as described in Section 5.4.3) it went unnoticed by many children. Therefore, we conclude that a search engine is more supportive for children's search performance than a menu structure. However, we conclude this on the condition that children already have experience using a search engine and that the search engine gets noticed by children by placing it on a location that is expected by children.

We falsified our hypothesis that the layout of the three types of website design was of influence on children's search performance, based on the results of our experiment. Bruckman (2008) reported that children enjoy a playful approach, which means that

graphics should be used that children can relate to. We also experienced that children enjoyed a playful approach, but mainly when the playful approach was realized in the visual design of the interface. When the navigation design of the interface was playful, the children were far less positive about the interface, because this playful navigation approach lowered usability, as was the case on the Image map website type in our research.

Most children preferred a more conventional menu (i.e. Classic) over the menus incorporated in graphical metaphors, as was the case on the Image map website type. Also, both the children in the explorative study and in the experimental study showed that it worked well for them when the textual categories were accompanied by a picture as was the case with the Classical play websites in both studies.

6.3.4 Interaction dimension

Search experience from a performance perspective

The children in our research did not experience many problems using a search engine, as was reported in previous research (Borgman et al., 1995; Schacter et al., 1998 and McCrory, Wallace et al., 2000). Correct spelling and formulating a single keyword instead of natural language questions were difficulties that children experienced, but in general, children could quite easily work with the search engine. Most likely, the reason for this is that children nowadays have more experience using search engines than ten years ago (Druin et al., 2009).

Children experienced more problems with navigation on the websites in our studies. Just as reported by Tanni (2008), the children in our studies also often demonstrated 'fast surfing'; they often searched information for the 'right' answer instead of constructing personal understanding. This often resulted in children being easily satisfied with an incorrect answer that seemed relevant to them, without really thinking it through. The children in our studies were often reactive searchers, as was also reported by Chen (2003). Most children did not use sophisticated analytical search strategies and they often repeated unsuccessful searches by returning to the same incorrect subcategories or search results. This loopy style was also reported by Hirsh (1999). Some children were chaotic searchers. They made many moves within one search task, often backtracked and often deviated from a designated target, as was also reported by Chen (2003). Also, children's tendency to go back to the home page often and to stay very close to home - as was reported by McCrory Wallace *et al.* (2000) - was often the case in our studies.

Concerning information collection and evaluation, we recognized from prior research results that children mainly scan pictures and first paragraphs of web pages. They rarely read web pages to understand in depth, as reported by Walraven *et al.* (2008) in a literature study about information search behaviour of children and teenagers.

They had little patience to read a text to find the right answer to their search task. We also recognized prior research results about relevance criteria. Walraven *et al.* (2008) reported the following relevance criteria that are not considered by children: authority, recency, truthfulness, accuracy and validity. The children in our studies also did not consider these criteria to judge relevance of search results.

Search experience from an evaluation perspective

Besides pragmatic search performance, we were also interested in the affective side of children's experience while interacting with a search interface. Therefore, following the examples of Hassenzahl (2004), Van Schaik and Ling (2008) and Tuch *et al.* (2012) we analyzed children's emotional reactions during information search and measured feelings, perceptions of instrumental and non-instrumental qualities and evaluations of beauty, goodness and fun using subjective survey methods. We also measured emotional arousal using an objective research method.

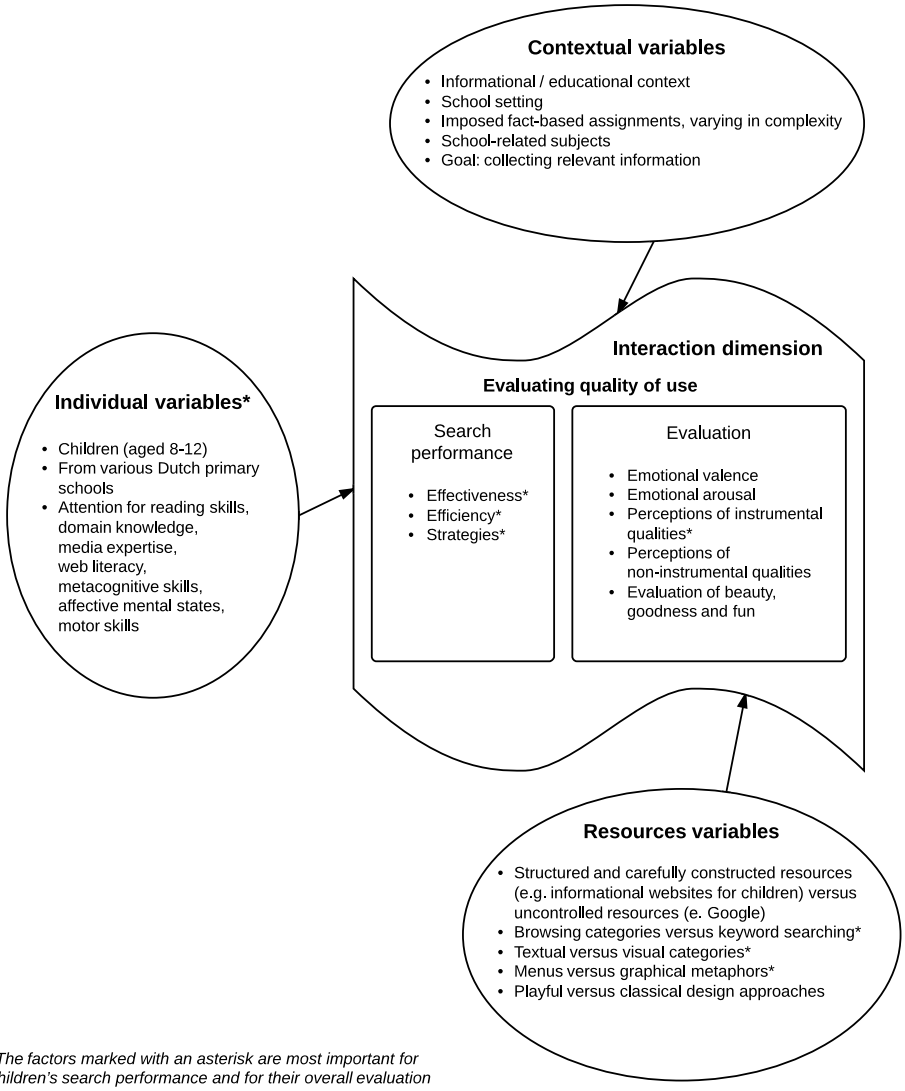
We experienced a lot of problems with the measuring of the affective side of user experience with children. Many children tended to satisfy the test instructor by giving positive answers, probably because they thought these were socially acceptable. For example, children had a large tendency to indicate the highest, most positive scores on the scales when scales were used to elicit opinions about the websites (as reported by Markopoulos et al., 2008). Furthermore, our transformation of the AttracDiff 2 questionnaire (as composed by Hassenzahl et al., 2003) from verbal to picture anchors for the semantic differential items, proved to be quite disappointing. Although the reliability of the survey items was very high, the items were not valid at all. It turned out that the children interpreted the items entirely differently than we had intended. Children were also not capable of making a distinction between instrumental and non-instrumental qualities of the websites. They tended to turn all items into the poles 'positive versus negative', as if all items were scales to measure valence.

The same problem was experienced with the evaluation of the interfaces for beauty, goodness and fun. It turned out that some of the children found it difficult to make a clear distinction between these constructs. These constructs appeared quite similar to the children, as was also reported by Read *et al.* (2002). This difficulty of using survey questions with children and the issues with the validity and reliability of survey answers given by children was also reported by Horton *et al.* (2011). Fortunately, none of the children indicated having problems with the meaning of these items (Section 5.5.1), from which we believe that most children did interpret these constructs as was intended by us.

In conclusion, we learned that studying the affective side of user experience with children is not the same as studying this with adults. Researchers should be very careful

in using research tools that were initially developed for research with adults, to measure User Experience components with children.

However, we learned from the User Experience components that we were able to measure validly, that the affective side of user experience is most likely less important for children than it is for adults. Children do appreciate non-instrumental qualities, such as a playful approach on a search interface. However, usability proved to be much more important for children for their overall appraisal of a search interface. There-



**The factors marked with an asterisk are most important for children's search performance and for their overall evaluation of a search interface according to our research.*

Figure 6.8: Factors of information interaction in our research

fore, we recommend designers to be modest with a playful approach when it comes to informational search interfaces for children. Children appreciate playful visual design characteristics, but when it comes to pragmatic quality, children appreciate and benefit more from classic navigation design than playful navigation design.

To summarize this section, we present Figure 2.13 once more (see Figure 6.8 below), but now we have marked the variables (with an asterisk) that proved to be most important for children's information search behaviour, based on our research results.

6.4 Reflection on the research methods used

The results of our research that are reported in this dissertation were all measured using methods that we checked for reliability and validity. In our final experiment, however, some of the findings could not be reported, because the methods used for these measures turned out to be invalid in our research with children. Still, even from the use of invalid methods, lessons can be learned for future research with children. In this section, such lessons will be discussed.

6.4.1 Measuring search behaviour and performance

Logging the activities of the children and measuring pragmatic variables proved to be an effective method in the studies. The method was unobtrusive for the children and could run alongside the other evaluation methods, as reported by Markopoulos *et al.* (2008).

Using the eye-tracker was a valuable method for the explorative study, because children's screen activities and spoken comments could be very precisely combined with the area on the screen on which the children focused their attention. For the purpose of the explorative study to explore children's search behaviour, it was very useful to study children's eye movements over the screen. However, the purpose of the experiment was not to explore children's search behaviour in detail, but to measure quantitative data of children's search performance (i.e. required clicks and time). For the experiment, logging screen events and recording screen activities in combination with children's spoken comments (i.e. voluntary think-aloud method) without their eye movements provided enough insight into children's search performance (Elling et al., 2012).

In the explorative study, we conducted post-task interviews after each search task instead of using think-aloud protocols, to prevent extra cognitive load on the children during the search process itself. However, the results of the post-task interviews did not add value to the observations made by logging the screen activities and making video and audio recordings of the children during search performance. Children did not report many problems or successes and had trouble reflecting on their own search process, as was also reported by Baauw and Markopoulos (2004). Therefore, we did not

conduct post-task interviews in the experiment. The observations made by logging the screen activities and making video and audio recordings to record spontaneous verbal comments were sufficient to reach our research goals.

It might be that more value was added to the observations made by logging the screen activities, by using a *retrospective* think-aloud method, in which participants work in silence and verbalize their thoughts afterwards while watching a recording of their performance. Elling, Lentz and De Jong (2011) used two types of retrospective think-aloud methods: one in which participants watched a recording of their performance with the corresponding dynamic gaze trail of their eye movements and one without their eye movements. Eye movements might help participants remembering in more detail what they did or what they thought during task performance. Elling *et al.* (2011) did not find additional value in showing participants their eye movements: the output of the two variants they compared did not differ. Future research with children should be conducted to study whether a retrospective think-aloud method adds value in research with children. However, in line with Elling *et al.* (2011), we believe that showing eye movements will probably not add value in research with children, as it does not add value in research with adults. It will most likely be distracting rather than helpful for children in explaining their own search behaviour to the test instructor. Especially when considering the fact that for children it is difficult to reflect on their own search behaviour, because their metacognitive skills – required for this type of self-monitoring - are not fully developed yet (Brown & DeLoache, 1978).

6.4.2 Smiley scales: measuring difficulty scores and emotional valence and arousal

We used smiley scales in the explorative study to measure children's difficulty scores on the search tasks conducted (see Figure 4.6). We did not find a clear tendency of children to give extreme positive ratings as Read *et al.* (2002) did. This might be caused by the fact that children were asked about how difficult the task was (see 4.4.3), so not only the interface was evaluated, but also the interaction with this interface. In other words, children could also blame themselves for evaluating the task as 'difficult' instead of blaming the system as being difficult, which may reduce the tendency to give extreme positive ratings. Children also did not report problems with the neutral position on the scale which had a face with as straight-line mouth, as Read (2008) found. Moreover, there was clearly a tendency to choose this neutral position on the smiley-scale, which had the label 'normal'. Apparently, children were not very agitated by the tasks and interfaces and they often evaluated the tasks and interfaces as "normal; not very easy, but also not very difficult".

We used a similar smiley scale in the experiment by using the SAM-scale (Lang, 1980) for measuring children's emotional valence. As this smiley scale was not about difficulty

of conducting the search task, but about children's feelings with the interface, we did find a tendency for children to give extreme positive ratings, just as Read *et al.* (2002) did. Children could not really blame themselves when evaluating the system compared to evaluating the difficulty of the task in the explorative study. Therefore, they were more prone to give positive ratings to the system. The SAM-scale for emotional valence proved to be valid with children, because the children in our research did interpret the smiley scale used for valence as intended; as a negative versus a positive feeling towards the search interface.

We also used the SAM-scale (Lang, 1980) for measuring children's emotional arousal. Unlike Greenbaum *et al.* (1990), we could not prove this SAM-scale to be a valid method in research with children, because the children in our research did not interpret the arousal scales as low versus high aroused feelings as was intended. These results might have been caused by the different settings in which the method was used. The pictures in the SAM-scale for emotional arousal might have been related easier to 'fear for the dentist' by children than the feeling experienced with using a website.

6.4.3 Measuring perceptions of pragmatic and hedonic quality

Our version of the AttracDiff 2 questionnaire turned out to be invalid in research with children, because children did not give the same meaning to the semantic differential scales as was intended. Almost all picture anchors for the semantic differential scales for the pragmatic and hedonic items turned out to be invalid for use with children in our experiment. Children did not interpret the picture anchors as intended by the developers of the AttracDiff 2 questionnaire (Hassenzahl et al., 2003). The children only made a clear distinction between which end of the scale was intended as positive and which end of the scale as negative in their opinion. The instrument did not make clear to the children what the difference is between perceptions of hedonic and pragmatic quality of the websites. The pragmatic and hedonic items were all seen as the same construct of valence (i.e. positive versus negative). Although we did not test the verbal semantic differential scales as used by Hassenzahl *et al.* (2003) with the children, we do not think that children will understand the difference between perceptions of hedonic and pragmatic quality when verbal instead of picture anchors are used for the semantic differential scales.

6.4.4 Observing affective responses

A more objective method to measure children's emotional experience than asking about their feelings is to observe their expressive reactions towards the websites, as we did in our explorative study and in a subset of the data collected in the experiment. By observing children's spontaneous utterances (termed "voluntary think-aloud" by Donker & Reitsma, 2004), we hoped to get a clear picture of the children's product experience. This

method proved to be very helpful in our qualitative comparison of the children's product experience between the three different websites in the explorative study. Some children did spontaneously utter emotional reactions towards the websites. It turned out that most expressive reactions were attributed to usability (pragmatic quality) of the websites and not to non-instrumental qualities, such as aesthetics and playful design. We saw the same pattern in the analysis of a small sample of the recorded data in the experiment. Some children did spontaneously utter emotional reactions (both positive and negative) that were almost all attributed to usability issues, just as in the explorative study.

6.4.5 Measuring evaluation scores

In the explorative study, we asked children to rank the websites that they found best and that they found worst for future use and everything in between. They had no trouble ranking the best and the worst websites but they had much more trouble ranking the websites 'in between' the best and the worst. Read *et al.* (2002) also reported this problem that children could easily rank the items that they found the best or the worst, but that it was much more difficult to assign items to the ranks 'in between'. These ranks were often not assigned at all. Therefore, less value should be attached to the ranking of the websites that were ranked 'in between'. In the experiment, we asked children to give the interfaces report marks for beauty, goodness and fun as evaluation scores. Although we did find differences in scores for the evaluations, we doubt whether children were able to make a clear distinction between the concepts beauty, goodness and fun, based on some of the comments made by the children, as was also reported by Read *et al.* (2002).

6.4.6 Measuring physical emotional arousal

Finally, we used an objective method to measure emotional arousal by using the Q Sensor. This method turned out to be valid to measure emotional arousal with children in contrast to the subjective methods used in our experiment, because we found a significant order effect on the amount of electrodermal activity (i.e. a physical property for emotional arousal). Children that watched the film after the search task performance showed a higher number of electrodermal responses per minute during the task performance than children that watched the film before the search task performance. This proved to us that we were actually measuring emotional arousal by using the Q Sensor. However, we experienced that the type of information tasks given to the children are not suited for use of the Q Sensor in general. With many children, we did not measure any electrodermal activity at all, while the other physical properties, electrode temperature and acceleration did show activity, which proves that it was not caused by malfunctioning of the device. Apparently, searching information on a computer does not cause a lot of emotional arousal with children, which we think is an interesting result. Differences

between several types of search interfaces are too small to measure differences in emotional arousal.

Based on our research, parts of the User Experience models presented in research with adults (Hassenzahl, 2004; Thüring & Mahlke, 2007; Van Schaik & Ling, 2008) turned out to be not the same in research with children. Although we could not entirely reproduce User Experience research concerning perceptions of pragmatic and hedonic quality in our experiment, we did show the difficulty of reproducing User Experience research methods with children. More research is needed to develop valid methods to measure children's perceptions of pragmatic and hedonic quality and to provide more evidence that usability has a stronger influence on children's evaluation of an interface than visual aesthetics, stimulation and identification.

6.5 Recommendations for designers and educators

The first step in our research project was to identify design conventions used in current informational websites for children. The analysis of the design characteristics of the 100 informational websites led to a categorization of website types, with an outcome of three main website types. The next steps in our research project were mainly focused on the overall implications of these three different types of web design on children's search performance. When considering the implications of the specific design conventions of these three types of website design (identified in Chapter 3) on children's search performance that we observed within our studies (in Chapter 4 and 5), we recommend the following do's and don'ts for children's informational interface design (see Table 6.1), which we will discuss in more detail below.

Table 6.1: Do's and don'ts for children's informational websites based on the observations made in our studies

	Do's	Don'ts
Visual design	· The main menu placed at a conventional location (at top or left of the screen) · The search box placed at (top) right of the screen · Ordered pages (no clutter) · Solid background colors · Use of graphics · Use of animations · Use of a focal point · Consistency in layout	· Use of non-solid backgrounds · Use of playful fonts · Little use of graphics · Little use of animations · Low screen density
Navigation design	· Presentation of clear (non-ambiguous) main categories · Presentation of link labels with text and icons · Consistency in the menu structure	· Use of Image map navigation · Presentation of only-text link labels
Functionality	· Providing a search engine · Providing a separate home button	· Use of pop-up windows or 'layered' screen elements

6.5.1 Visual design

The children in our research did appreciate visual design elements, such as a beautiful or fun focal point (e.g. the dog with glasses in Figure 6.9), colors, graphics and animations. Terms that they used for this type of visual design were 'nice' and 'funny'. Pages with low screen density and a lot of whitespace were less attractive for children. However, when a page was too crowded or cluttered to be transparent, the children appreciated the interface less. Also, use of non-solid backgrounds and playful fonts were appreciated less, because this caused the interfaces to be less transparent and therefore, less user-friendly. Therefore, the basis of the interface should be ordered, but 'dressed up' with 'nice' and 'funny' elements. This basic design should be clear and consistent throughout all pages of a website and the 'nice' and 'funny' elements should not go at the cost of order.

Important components of an informational interface, such as a menu or a search engine should be placed at a conventional location, because children expect these components at such conventional locations. For example, the menu can best be placed horizontally at the top or vertically at the left of the screen and the search engine can best be placed at the (top) right of the screen. Because of technical restrictions we were not able to place the search engine in our experimental stimuli at the right hand side of the screen. The reason for this was that the publisher of the 'Junior Winkler Prins online encyclopedia' was responsible for the technical implementation of the experimental websites. The original interface (i.e. the Image map version) was built in a Flash-environment (by a design company) which the publisher could not adjust. We suspect that this was an important reason that many children did not see – and therefore did not use – the search engine on these experimental interfaces. However, further research should provide more evidence for this hypothesis.

6.5.2 Navigation design

The children in our research appreciated a conventional (i.e. classic) navigation design. We did not find evidence that children's search performance was better using a classic navigation design in comparison to an Image map navigation design. However, children did experience a lot of problems and frustrations when using the Image map menus in our studies and they appreciated the Image map interfaces the least of the interfaces used in our studies.

When the navigation design of the interface was playful, the children were far less positive about the interface, because this playful navigation approach lowered usability, as was the case on the Image map website type in our research. As opposed to Meloncon *et al.* (2010) and Nielsen and Gilutz (2002) we do not recommend using a visual map for navigation in an informational search interface for children.

Also, category labels that were vague or abstract were not appreciated by the children. In our explorative study, children could often not imagine the type of information that could be found by clicking on a category label. We recommend the use of concrete, factual, scholarly terms such as 'health', 'arts', 'sports', 'nature', 'science', etcetera. It is also important for children that this menu structure is consistent throughout a website, to provide constant guidance throughout the search process.

Another important lesson learned from our research project, is that the combination of textual category labels with a relevant picture or icon works best for children (see the main menu horizontally at the top of Figure 6.9). The children in our studies were not always sure about their reading skills. Confirmation about the meaning of a textual label in the form of a relevant icon or picture was therefore appreciated by the children. Furthermore, some children are more visually than textually oriented and these children often only scan pictorial information instead of textual information in search of a relevant category (as described in Chapter 4). Therefore, we recommend adding graphics to textual categories to support children's visual scanning of categories, as was also recommended by Large and Beheshti (2005).

6.5.3 Functionality

In our explorative study, Google did not always guide children (in a safe manner) to relevant information. Therefore, based on our own research findings, we recommend structured and carefully constructed resources (i.e. informational websites especially designed for children) for children's information search tasks. As opposed to Meloncon *et al.* (2010), we definitely recommend offering a search engine on an informational search interface for children. Children's search performance in both our explorative study and our experiment was much more effective and efficient when using a search engine than when using the menu structure on these interfaces. We did not study the implications of the use of a search engine for children's overall learning. However, we can imagine that children's overall learning is supported more by following a hierarchical menu structure than by using 'a black magic box', because that is what a search engine actually is. Future research should study such implications of search engine use for children's overall learning.

6.5.4 Our recommendation: a Classical play design approach

In our studies, a Classical play design approach proved to be the best of both worlds for children: a classic navigation approach accompanied by playful ('funny' and 'nice') visual design characteristics, as can be seen in Figure 6.3. However, the availability of a search engine (preferably at the right side of the screen) is maybe even the most important

characteristic for a children's search interface, to best support children in effective and efficient information search.

6.5.5 Recommendations for educators

The focus of our research project was on analysis and adaptation of search interfaces and we did not focus on adaptation of children's behaviour and web literacy. We believe that an interface must be designed in such a manner that it supports children in effective information search, which was our major research objective. Of course, in our research, we also experienced the differences between children's information search performance, caused by individual differences, such as differences in domain knowledge, media experience, reading abilities and web literacy. Therefore, we believe that besides supportive interface design, educators play a major role in children's search behaviour and performance. Children's reading abilities, for example, have influence on the processing of a menu, the selection of a relevant category and on the decision whether a found information page is relevant for the search task. Also, web literacy and media experience have influence on children's search performance. Children who know how to use a menu or a search engine experience far fewer problems than children who do not know the basics of searching or browsing. Furthermore, metacognitive skills are of influence on children's search performance. Children that monitor the process and stay focused on their target are much more efficient information searchers than children who do not monitor the process and get distracted by every irrelevant element that they cross along the way. Therefore, teaching children how to formulate an effective search query, how to browse main categories, how to select relevant link labels, how to scan information pages for relevant information and how to monitor the search process, are very important elements for educators in supporting children's search performance. For research on children's web literacy, we think that the work of Els Kuiper, reported in her dissertation titled "Teaching Web literacy in primary education" (2007), is of much value.

6.6 Implications for current design trends and recommendations for future research

Technology is constantly evolving. When we started this project, iPads did not exist, and now, for many children it is almost impossible to imagine leisure without the iPad. Orthopedic surgeons even warn against children developing hunched backs, because they spend so much time hunched over their smart phones or tablets (Volkskrant, 07/08/2013). Our research is aimed at technology that belongs to the desktop paradigm, which is still daily practice at most Dutch schools concerning devices for digital education anno 2013.

The desktop computers that have been used for years now with the mouse and keyboard as input devices are not very intuitive for children. These computers were not designed with the interactions of children in mind or with the goal of supporting learning (Inkpen, 2001). Touch screen technology is a more natural input device than the traditional options. The most important advantage of touch screens is that the input device is also the output device. There is a sense of immersion, because the user is able to touch, feel and manipulate objects on a screen (Holzinger, 2003). Therefore, this might be a more effective input device for children. Future research should give more insight into the differences between interacting with a traditional desktop computer interface and interacting with a touch-screen interface for children's search behaviour. Within a few years from now, children's desks at school will most likely be touch tables through which all school work can be conducted. Or maybe all children will wear 'Google-glass'-type devices through which all school work can be conducted. Who knows?

Independent of the technological developments and the different ways of interacting with devices that are evolving constantly, we do think that our research findings will remain valuable and applicable to future technological solutions. Whether children use a mouse, a touch device or speak to a device, principles of how information should be organized and designed - so that children can reach their goals - will mostly stay the same. When the goal is to find relevant information, effective and efficient information search is more important for children than the fun experience of interacting with a beautiful, playful search interface.

At the beginning of this dissertation, we hypothesized that the Google-type search engine is not the best search interface for children's search behaviour. Children's cognitive skills are still developing and they are not yet able to judge whether information is relevant for their search task or whether the source of information is trustworthy. Also, the amount of information that is provided by Google-type search engines would be overwhelming for children. We did find some proof to support this hypothesis, as described in paragraph 6.2.3., but we did not entirely validate this hypothesis and future research is needed to provide more evidence.

We do acknowledge that the use of a search engine is more effective and efficient for children's search performance than browsing categories. We also acknowledge that the tools provided by Google-type search engines, such as spelling corrections and query suggestions, are very effective for children's search performance. Children were not confronted with harmful content in our study, because we controlled the type of information that they searched for. However, earlier studies (Wishart, 2004; Valcke, De Wever, Van Keer & Schellens, 2011) learned that search engines for children should be based on structured and carefully constructed resources and be equipped with powerful search filters to protect children from harmful content. In the field of 'information retrieval' research is conducted on how search systems can filter and find information that is

suitable for children (Eickhoff, Serdyukov & deVries, 2010; Elliot, Azzopardi, Glassey & Polajnar, 2010).

Furthermore, the emphasis on usability and fluency of information searching of the Google-type search engines is more important for children's search behaviour than non-instrumental qualities such as playful interaction experiences. Children do appreciate 'fun' and 'nice' visual design elements, such as images or animations, which are absent in the Google-type search engines. Also, because children are often visually oriented, the addition of images or icons to category link labels or to search result labels is likely to work well for children. Researchers even developed an interface that enabled searching through adaptable picture directories. Children could construct queries using pictures (Polajnar, Glassey, Gyllstrom & Azzopardi, 2011). This might be a good solution for children that have problems with regular keyword searching.

However, these 'playful' design elements should not be used at the cost of usability. Interfaces for children should therefore always be tested to evaluate the usability of the interface. Preferably, such a usability test should be conducted with children (as we did in our explorative study and experiment) and not only by using checklists with heuristics. Nobody knows better what does and does not work for children than the children themselves.

One thing has become clear from our research: designers of children's informational websites should pay more attention to the usability of search interfaces than to the playfulness of interface design to support children in effective information search.

ZOEK

Bibliography

Adkisson, H. (2002). *Identifying de-facto standards for E-commerce websites.* (Unpublished Master of Science Thesis). University of Washington, USA.

Ahtikari, J., & Eronen, S. (2004). On a journey towards web literacy - The electronic learning space netro (Unpublished doctoral dissertation). University of Jyvaskyla, Department of Languages, Jyväskylä, Finland.

Baauw, E., & Markopoulos, P. (2004). A comparison of think aloud and post-task interview. Paper presented at the 2004 Conference on Interaction Design and Children: Building a Community. Maryland, USA (pp. 115-116).

Barfield, L. (2004). *Design for new media*. England: Pearson Education Limited.

Bar-Ilan, J., and Belous, Y. (2007). Children as architects of web directories: An exploratory study. *Journal of the American Society for Information Science and Technology, 58*(6), 895-907.

Belkin, N. J. (1993). Interaction with texts: Information retrieval as information seeking behavior. *Information Retrieval, 93*, 55-66.

Bevan, N. (1995). Measuring usability as quality of use. *Software Quality Journal, 4*(2), 115-130.

Bilal, D. (2000). Children's use of the yahooligans web search engine: I. cognitive, physical, and affective behaviours on fact-based search tasks. *Journal of the American Society for Information Science, 51*(7), 646-665.

Bilal, D. (2001). Children's use of the yahooligans! web search engine: II. cognitive and physical behaviours on research tasks. *Journal of the American Society for Information Science and Technology, 52*(2), 118-136.

Bilal, D. (2002). Children's use of the yahooligans! web search engine. III. cognitive and physical behaviours on fully self-generated search tasks. *Journal of the American Society for Information Science and Technology, 53*(13), 1170-1183.

Bilal, D., & Kirby, J. (2002). Differences and similarities in information seeking: Children and adults as web users. *Information Processing and Management, 38*(5), 649-670.

Bilal, D., & Watson, J. S. (1998). Children's paperless projects: Inspiring research via the web. Paper presented at the 64th General Conference of the International Federation of Library Associations & Institutions, Amsterdam, the Netherlands.

Bilal, D., and Wang, P. (2005). Children's conceptual structures of science categories and the design of web directories. *Journal of the American Society for Information Science and Technology*, *56*(12), 1303-1313.

Borgman, C. L., Hirsh, S. G., Walter, V. A., and Gallagher, A. L. (1995). Children's searching behaviour on browsing and keyword online catalogs: The science library catalog project. *Journal of the American Society for Information Science, 46*(9), 663- 684.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behaviour Therapy and Experimental Psychiatry, 25*(1), 49-59.

Brown, A. L., and DeLoache, J. S. (1978). Skills, plans, and self-regulation. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 3-35). New Jersey: Hillsdale Erlbaum.

Bruckman, A., Bandlow, A., & Forte, A. (2008). HCI for kids. In J. Jacko, & A. Sears (Eds.), *Handbook of human-computer interaction* (pp. 793-809) Lawrence Erlbaum Associates.

Bruner, J. S. (1966). *Toward a theory of instruction.* Belknap Press.

Carey, S. (1987). *Conceptual change in childhood.* Bradford Books/MIT Press.

Carroll, J. M., & Thomas, J. C. (1988). Fun. *ACM SIGCHI Bulletin, 19*(3), 21-24.

Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology, 27*(1), 27-36.

Chen, C. (2006). *Information visualization: Beyond the horizon.* London, UK: Springer.

Chen, S. L. (2003). Searching the online catalog and the world wide web. *Journal of Educational Media & Library Sciences, 41*(1), 29-43.

Cooper, L. Z. (2002). Methodology for a project examining cognitive categories for library information in young children. *Journal of the American Society for Information Science and Technology, 53*(14), 1223-1231.

Creswell, J. W., & Clark, V. L. P. (2007). *Designing and conducting mixed methods research.* Wiley Online Library.

Croft, W. B. (1993). Knowledge-based and statistical approaches to text retrieval. *IEEE Expert, 8*(2), 8-12.

De Belder, J., Deschacht, K., & Moens, M. (2010). Lexical simplification. Paper presented at the ITEC2010: 1st International Conference on Interdisciplinary Research on Technology, Education and Communication, Kortrijk, Belgium.

Derr, R. L. (1983). A conceptual analysis of information need. *Information Processing & Management, 19*(5), 273-278.

Desmet, P. M. A. (2008). Product emotion. In H. N. J. Schifferstein, & P. Hekkert (Eds.), *Product experience* (pp. 379-398). San Diego: Elsevier.

Desmet, P., & Hekkert, P. (2007). Framework of product experience. *International Journal of Design, 1*(1), 57-66.

Detken, K., Martinez, C., & Schrader, A. (2009). The search wall: Tangible information searching for children in public libraries. Paper presented at the TEI '09: 3rd International Conference on Tangible and Embedded Interaction, Cambridge, United Kingdom (pp. 289-296).

Dillon, A. (2002). Beyond usability: Process, outcome and affect in human-computer interactions. *Canadian Journal of Library and Information Science, 26*(4), 57-69.

Donker, A., & Reitsma, P. (2004). Usability testing with young children. Paper presented at the 2004 Conference on Interaction Design and Children: Building a Community, Maryland, USA (pp. 43-48).

Dresang, E. T. (2005). The information seeking behavior of youth in the digital environment. *Library Trends, 54*(2), 178-196.

Druin, A. (2003). What children can teach us: Developing digital libraries for children with children. *The Library Quarterly, 75*(1), 20-41.

Druin, A. (2005). Introduction to "children's access and use of digital resources". *Library Trends, 54(2)*, 173-177.

Druin, A., Hutchinson, H., Foss, E., Hatley, L., Golub, E., Leigh Guha, M., & Fails, J. (2009). How children search the internet with keyword interfaces. Paper presented at the 8th International Conference on Interaction Design and Children, Como, Italy.

Duarte Torres, S., Hiemstra, D., & Serdyukov, P. (2010). An analysis of queries intended to search information for children. Paper presented at the Third Symposium on Information Interaction in Context, New Brunswick, USA (pp. 235-244).

Dunn, K. (2002). Assessing information literacy skills in the california state university: A progress report. *The Journal of Academic Librarianship, 28*(1–2), 26-35.

Eickhoff, C., Serdyukov, P., & de Vries, A. P. (2010). Web page classification on child suitability. Paper presented at the 19th ACM International Conference on Information and Knowledge Management, Toronto, Canada (pp. 1425-1428).

Elling, S., Lentz, L., & de Jong, M. (2011). Retrospective think-aloud method: Using eye movements as an extra cue for participants' verbalizations. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, Canada (pp. 1161-1170).

Elling, S., Lentz, L., & de Jong, M. (2012). Combining concurrent think-aloud protocols and eye-tracking observations: An analysis of verbalizations and silences. *IEEE Transactions on Professional Communication,* 55(3), 206-220.

Elliot, D., Glassey, R., Polajnar, T., & Azzopardi, L. (2010). Finding and filtering information for children. Paper presented at the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, Geneva, Switzerland (pp. 702).

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis; verbal reports as Data*. Cambridge, MA: Bradford Books/ MIT Press.

Farkas, D. K., and Farkas, J. B. (2000). Guidelines for designing web navigation. *Technical Communication, 47*(3), 341-358.

Fitzgerald, M. A. (2005). Skills for evaluating web-based information. Paper presented at the Internet Credibility Symposium, Seattle, USA.

Fu, W. T., & Pirolli, P. (2007). SNIF-ACT: A cognitive model of user navigation on the world wide web. *Human-Computer Interaction, 22*(4), 355-412.

Gossen, T., & Nürnberger, A. (2013). Specifics of information retrieval for young users: A survey. *Information Processing & Management, 49*(4), 739-756.

Gossen, T., Low, T., & Nürnberger, A. (2011). What are the real differences of children's and adults' web search. Paper presented at the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, Beijing, China (pp. 1115-1116).

Greenbaum, P. E., Turner, C., Cook 3rd., E. W., & Melamed, B. G. (1990). Dentists' voice control: Effects on children's disruptive and affective behaviour. *Health Psychology, 9*(5), 546-558.

Harbeck, J. D., and Sherman, T. M. (1999). Seven principles for designing developmentally appropriate websites for young children. *Educational Technology, 39*(4), 39-44.

Hartmann, J., Sutcliffe, A., & Angeli, A. D. (2008). Towards a theory of user judgment of aesthetics and user interface quality. *ACM Transactions on Computer-Human Interaction, 15*(4), 1-30.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction*, *19*(4), 319-349.

Hassenzahl, M. and Tractinsky, N. (2006). User experience–a research agenda. *Behaviour & Information Technology*, *25*(2), 91-97.

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualitat [AttracDiff: A questionnaire to measure perceived hedonic and pragmatic quality]. In J. Ziegler, & G. Szwillus (Eds.), Mensch&Computer: Interaktion in bewegung (pp. 187-196). Stuttgart, Leipzig: B.G. Teubner.

Hassenzahl, M., Platz, A., Burmester, M., & Lehner, K. (2000). Hedonic and ergonomic quality aspects determine a software's appeal. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, The Hague, The Netherlands (pp.201-208).

Hassenzahl, M., Schöbel, M., and Trautmann, T. (2008). How motivational orientation influences the evaluation and choice of hedonic and pragmatic interactive products: The role of regulatory focus. *Interacting with Computers, 20*(4-5), 473-479.

Herring, S. C., Scheidt, L. A., Bonus, S., & Wright, E. (2004). Bridging the gap: A genre analysis of weblogs. Paper presented at the 37th Annual Hawaii International Conference on System Sciences, Computer Society Press.

Hirsh, S. G. (1997). How do children find information on different types of tasks? Children's use of the science library catalog. *Library Trends, 45*(4), 725-745.

Hirsh, S. G. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science, 50*(14), 1265-1283.

Holzinger, A. (2003). Finger instead of mouse: Touch screens as a means of enhancing universal access. *Universal Access Theoretical Perspectives, Practice, and Experience Lecture Notes in Computer Science, 2615*, 387-397.

Horton, M., Read, J.C. and Sim, G. (2011). Making your mind up? The reliability of children's survey responses. In L. Little & L. Coventry (Eds.), *25th BCS Conference on Human-Computer Interaction* (pp. 437-438) Swinton, UK: British Computer Society.

Hourcade, J. P. P., Druin, A., Sherman, L., Bederson, B. B., Revelle, G., Campbell, D., Weinstein, B. (2002). SearchKids: A digital library interface for young children. Paper presented at the CHI '02: Extended Abstracts on Human Factors in Computing Systems, Minneapolis, Minnesota, USA (pp. 512-513).

Hutchinson, H. B., Bederson, B. B., & Druin, A. (2006). The evolution of the international children's digital library searching and browsing interface. Paper presented at the 2006 Conference on Interaction Design for Children, Tampere, Finland.

Hutchinson, H., Druin, A., Bederson, B. B., Reuter, K., Rose, A., and Weeks, A. C. (2005).

How do I find blue books about dogs? The errors and frustrations of young digital library users. Paper presented at the International Conference on Human-Computer Interaction (HCI), Las Vegas, Nevada, USA.

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

Inkpen, K. M. (2001). Drag-and-drop versus point-and-click mouse interaction styles for children. *ACM Transactions on Computer-Human Interaction, 8*(1), 1-33.

Irtel, H. (2007). PXLab: The psychological experiments laboratory (2.1.11 ed.). Mannheim: University of Mannheim.

ISO 9241-11 standard: 'Guidance on Usability' (1998

John W. Creswell. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches.* Sage Publications, Inc.

Jonassen, D. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*(4), 63-85.

Jones, S., and Degrow, D. (2011). Fortune 500 home pages design trends. *IEEE Transactions on Professional Communication*, *54*(1), 18-30.

Juvina, I., & Oostendorp, H. v. (2008). Modeling semantic and structural knowledge in web navigation. *Discourse Processes, 45*, 346-364.

Kano, A., Horton, M., & Read, J. C. (2010). Thumbs-up scale and frequency of use scale for use in self-reporting of children's computer experience. Paper presented at the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, Reykjavik, Iceland (pp. 699-702).

Kao, G. Y., Lei, P., & Sun, C. (2007). Thinking style impacts on web search strategies. *Computers in Human Behavior, 24*(4), 1330-1341.

Kari, J. (1998). Making sense of sense-making: From meta-theory to substantive theory in the context of paranormal information seeking. Paper presented at the Nordis-Net Workshop: (Meta)Theoretical Stands in Studying Library and Information Institutions: Individual, Organizational and Societal Aspects, Oslo, Norway.

Kim, J., & Moon, J. Y. (1998). Designing towards emotional usability in customer interfaces—trustworthiness of cyber-banking system interfaces. *Interacting with Computers, 10*(1), 1-29.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163-182.

Kitajima, M., Blackmon, M. H., & Polson, P. G. (2000). A comprehension-based model of web navigation and its application to web usability analysis. Paper presented at the 14th Annual Conference of the British HCI Group (HCI 2000), People and Computers XIV: Usability Or Else!, Sunderland, UK.

Koyani, S. J., Bailey, R. W., Nall, J. R., Allison, S., Mulligan, C., Bailey, K., and Tolson, M. (2006). *Research-based web design & usability guidelines*. Washington: U.S. General Services Administration.

Krug, S. (2006). *Don't make me think!: A common sense approach to web usability* (2nd ed.).

Berkeley: New Riders.

Kuhlthau, C. C. (2004). *Seeking meaning: A process approach to library and information services* (2nd ed.). Westport: Libraries Unlimited.

Kuiper, E. (2007). Teaching web literacy in primary education. Vrije Universiteit, Amsterdam, The Netherlands.

Laarni, J. (2004). Aesthetic and emotional evaluations of computer interfaces. Paper presented at the Nordi CHI 2004 Workshop: Aesthetic Approaches to human–computer Interaction, Aarhus, Denmark.

Lang, P. J. (1980). Behavioural treatment and bio-behavioural assessment: Computer applications. In J.B. Sidowski, J.H. Johnson, T. A. Williams (Eds.), *Technology in Mental Health Care Delivery Systems* (pp. 119-137). Norwood, NJ, USA: Ablex Publishing.

Large, A., & Beheshti, J. (2000). The web as a classroom resource: Reactions from the users. *Journal of the American Society for Information Science, 51*(12), 1069-1080.

Large, A., & Beheshti, J. (2005). Interface design, web portals, and children. *Library Trends, 54*(2), 318-342.

Large, A., Beheshti, J., & Moukdad, H. (1999). Information seeking on the web: Navigational skills of grade-six primary school students. Paper presented at the ASIS Annual Meeting, Washington DC, USA.

Large, A., Beheshti, J., and Rahman, T. (2002). Design criteria for children's web portals:

The users speak out. *Journal of the American Society for Information Science and Technology, 53*(2), 79-94.

Large, A., Beheshti, J., Tabatabaei, N., & Nesset, V. (2009). Developing a visual taxonomy: Children's views on aesthetics. *Journal of the American Society for Information Science and Technology, 60*(9), 1808-1822.

Large, A., Nesset, V., Beheshti, J., & Bowler, L. (2006). "Bonded design": A novel approach to the design of new technologies. *Library and Information Science Research, 28*(1), 64-82.

Lavie, T., and Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of websites. *International Journal of Human-Computer Studies, 60*(3), 269-298.

Lazar, J. (2006). *Web usability - a user-centered design approach*. Boston, MA: Addison-Wesley.

Lazonder, A. W., & Rouet, J. (2008). Information problem solving instruction: Some cognitive and metacognitive issues. *Computers in Human Behavior, 24*, 753–765.

Lindgaard, G., & Dudek, C. (2003). What is this evasive beast we call user satisfaction? *Interacting with Computers, 15*(3), 429-452.

Logan, R. J. (1994). Behavioral and emotional usability: Thomson Consumer Electronics. In M. E. Wiklund (Ed.), *Usability in practice* (pp. 59-82). San Diego, CA, USA: Academic Press Professional, Inc.

Malone, T. W. (1980). *What makes things fun to learn? A study of intrinsically motivating computer games.* Technical Report No. CIS-7 (SSL-80-11). Palo Alto, California: Xerox Palo Alto Research Center.

Malone, T. W. (1984). Heuristics for Designing Enjoyable User Interfaces. Lessons from Computer Games. In J.C. Thomas, & M.L. Schneider (Eds.), *Human factors in computer systems* (pp. 1-12). Norwood, NJ, USA: Ablex Publishing.

Manning, C. D., Raghavan, P., & Schutze, H. (2008). *An introduction to information retrieval.* London: Cambridge University Press.

Marchionini, G. (1989). Information seeking strategies of novices using a full-text electronic encyclopedia. *Journal of the American Society for Information Science, 40*(1), 54-66.

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

Markopoulos, P., Read, J., MacFarlane, S., & Hoysniemi, J. (2008). *Evaluating children's interactive products.* Burlington: Elsevier.

McCrory Wallace, R., Kupperman, J., & Krajcik, J. (2000). Science on the web: Students online in a sixth-grade classroom. *The Journal of the Learning Sciences, 9*(1), 75-104.

Meloncon, L., Haynes, E., Varelmann, M., & Groh, L. (2010). Building a playground: General guidelines for creating educational Websites for children. *Technical Communication, 57*(4), 398-415.

Naidu, S. (2005). Evaluating the usability of educational websites for children. *Usability News*, 7(2).

Nastasi, B. K., Hitchcock, J. H., & Brown, L. M. (2010). An inclusive framework for conceptualizing mixed methods design typologies: Moving toward fully integrated synergistic research models. In A. Tashakkori & C. Teddlie (Eds.), *Sage handbook of mixed methods in social and behavioral research* (pp. 305-338). Thousand Oaks, CA: SAGE.

Nielsen, J., & Gilutz, S. (2002). *Usability of websites for children: 70 design guidelines based on usability studies with kids*. Nielsen Norman Group.

Nielsen, J., and Loranger, H. (2006). *Prioritizing web usability.* Berkely, CA, USA: New Riders Publishing.

Nielsen, J., and Tahir, M. (2002). *Home page usability - 50 websites deconstructed*.

Indianapolis, USA: New Riders Publishing.

Noordzij, M., Scholten, P., & Laroy-Noordzij, M. (2012). Measuring electrodermal activity of both individuals with severe mental disabilities and their caretakers during episodes of challenging behaviour. Presented at the 8th International Conference on Methods and Techniques in Behavioural Research: Measuring Behaviour, Utrecht, The Netherlands.

Piaget, J. (1964). Part I: Cognitive development in children: Piaget development and learning.

Journal of Research in Science Teaching, 2(3), 176-186.

Picard, R. W., & Klein, J. (2002). Computers that recognize and respond to user emotion: Theoretical and practical implications. *Interacting with Computers, 14*(2), 141-169.

Pijpers, R., Marteijn, T., Bosman, M., Berg, v. d.,W., & Dijkerman, E. (2008). *Klik en klaar. Een onderzoek naar surfgedrag en usability bij kinderen.* (No. 111). Den Haag: Stichting Mijn Kind Online.

Poh, M., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Transactions on Biomedical Engineering, 57*(5), 1243-1252.

Read, J. C. (2008). Validating the fun toolkit: An instrument for measuring children's opinions of technology. *Cognition, Technology & Work, 10*(2), 119-128.

Read, J. C., & MacFarlane, S. (2006). Using the fun toolkit and other survey methods to gather opinions in child computer interaction. Paper presented at the International Conference on Interaction Design and Children, Tampere, Finland.

Read, J. C., MacFarlane, S. J., & Casey, C. (2002). Endurability, engagement and expectations: Measuring children's fun. Paper presented at the International Conference on Interaction Design and Children, Eindhoven, The Netherlands.

Read, J.C. (2008). Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. *Cognition, Technology & Work, 10*(2), 119-128.

Revelle, G., Druin, A., Platner, M., Bederson, B. B., J.P., H., & Sherman, L. (2002). A visual search tool for early elementary science students. *Journal of Science Education and Technology, 11*(1), 49-57.

Rose, M., Rose, G. M., & Blodgett, J. G. (2009). The effects of interface design and age on children's information processing of web sites. *Psychology & Marketing, 26*(1), 1-21.

Rowlands, I. (2008). *Information behaviour of the researcher of the future.* London: British Library, JISC.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Schacter, J., Chung, G. K. W. K., & Dorr, A. (1998). Children's Internet searching on complex problems: Performance and process analyses. *Journal of the American Society for Information Science, 49*(9), 840-849.

Schaik, van, P., and Ling, J. (2008). Modelling user experience with websites: Usability,

hedonic value, beauty and goodness. *Interacting with Computers, 20*(3), 419-432.

Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer, & P. Ekman (Eds.), *Approaches to emotion* (pp. 293-317). Lawrence Erlbaum Associates.

Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal process in emotion: Theory, methods, research* (pp. 92-120). New York: Oxford University Press.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*(4), 695-729.

Serenko, A. (2007). Are interface agents scapegoats? Attributions of responsibility in human-agent interaction. *Interacting with Computers, 19*(2), 293-303.

Siegler, R. S. (1991). *Children's thinking* (2nd ed.). New Yersey: Prentice-Hall, Inc.

Sim, G., MacFarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education, 46*(3), 235-248.

Snijders, T., & Bosker, R. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2nd ed.). London etc.: Sage Publishers.

Spyridakis, J. H. (2000). Guidelines for authoring comprehensible web pages and evaluating their success. *Technical Communication, 47*(3), 359-382.

Strauss, A. L., and Corbin, J. (1990). *Basics of qualitative research - techniques and procedures for developing grounded theory.* Newbury Park, CA: Sage Publications, Inc.

Sweller, J., Van Merrienboer, J. J. G., & Paas, G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Tanni, M., & Sormunen, E. (2008). A critical review of research on information behavior in assigned learning tasks. *Journal of Documentation, 64*(6), 893.

Thuring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human-technology interaction. *International Journal of Psychology, 42*(4), 253-264.

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers, 13*(2), 127-145.

Tuch, A. N., Roth, S. P., Hornbæk, K., Opwis, K., & Bargas-Avila, J. A. (2012). Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Computers in Human Behaviour, 28*(5), 1596-1607.

Valcke, M., De Wever, B., Van Keer, H., & Schellens, T. (2011). Long-term study of safe Internet use of young children. *Computers & Education, 57*(1), 1292-1305.

Van der Sluis, F., & Van den Broek, Egon L. (2010). Using complexity measures in information retrieval. Paper presented at the Third Symposium on Information Interaction in Context, New Brunswick, USA (pp.383-388).

Van der Sluis, F., & Van Dijk, E. M. A. G. (2010). A closer look at children's information retrieval usage: Towards child-centered relevance. Paper presented at the Workshop on Accessible Search Systems held at the 33st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2010), Geneva, Switzerland (pp. 3-10).

Van Deursen, A. J. A. M., & Van Dijk, J. A. G. M. (2009). Using the internet: Skill related problems in users' online behavior. *Interacting with Computers, 21*(5-6), 393-402.

Van Rijsbergen, C. J. V. (1979). *Information retrieval* (2nd ed.). Newton, MA, USA: Butterworth-Heinemann.

Vygotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT Press.

Walraven, A., Brand-gruwel, S., & Boshuizen, H. P. A. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. *Computers in Human Behavior, 24*(3), 623–648.

Williams, T. R. (2000). Guidelines for designing and evaluating the display of information on the web. *Technical Communication, 47*(3), 383-396.

Wishart, J. (2004). Internet safety in emerging educational contexts. *Computers and Education, 43*(1-2 SPEC ISS.), 193-204.

Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management, 31*(2), 179-188.

Xu, D., Read, J., Mazzone, E., MacFarlane, S., & Brown, M. (2007). Evaluation of tangible user interfaces (TUIs) for and with children - methods and challenges. *Human-Computer Interaction. Interaction Platforms and Techniques. Lecture Notes in Computer Science, 4551*, 1008-1017.

Yarosh, S., Radu, I., Hunter, S., & Rosenbaum, E. (2011). Examining values: An analysis of nine years of IDC research. Paper presented at the 10th International Conference on Interaction Design and Children, Ann Arbor, Michigan.

Yusoff, Y. M., Ruthven, I., & Landoni, M. (2011). The fun semantic differential scales. Paper presented at the 10th International Conference on Interaction Design and Children, Ann Arbor, Michigan, USA (pp.221-224).

Zhang, P., & Li, N. (2005). The importance of affective quality. *Communications of the ACM, 48*(9), 105-108.

Web references

Chapter 3

Nielsen Wire. (2009). Growing up, and growing fast: Kids 2–11 spending more time online. The Nielsen Company. Retrieved July 19, 2011, from http://blog.nielsen.com/nielsenwire/online_mobile/growing-up-and-growing-fast-kids-2-11-spending-more-time-online. Nielsen, J. (2004). *The need for web design standards.* Nielsen Norman Group. http://www.nngroup.com/articles/the-need-for-web-design-standards.

Chapter 4

URL Classic website: http://schoolbieb.nl, last accessed at November 27, 2012. URL Classical play website: http://willemwever.nl, last accessed at November 27, 2012. URL Image map website: http://kids.kennisnet.nl/flash/#/thema/doelgroep, last accessed at November 27, 2012.

Chapter 6

Google's search market share from StatCounter Global Stats, http://royal.pingdom.com/2010/10/19/how-google-dominates-the-web, last accessed at September 22, 2013.

ZOEK

Appendices

Appendix A: Coding scheme for the corpus analysis (chapter 3)

1. Visual design

1a. Location or presence of items on the web page

(0 = not available, 1 = top left, 2 = top center, 3 = top right, 4 = horizontally on top of page, 5 =left side of page, 6 = center of page, 7 = bottom of page, 8 = right side of page)

Location of the logo

Location of the main menu

Location of the sub menu (only evaluated on the sub pages)

Location of the search window

Location of the servicelinks (e.g. about us, contact, privacy statement, help, etc.)

Presence and location of advertisement

Presence of shortcuts

Presence of interactivity (e.g. polls, guestbook, Q&A's, forum)

Location of home button

1b. Scanability

How many screens can the user scroll down?

(1 = 1-2 screens, 2 = 2-3 screens, 3 = 3 screen or more)

Are there scrolling pages in a smaller window within the page? (1 = yes, 0 = no)

Is there need to horizontally scroll on the page? (1 = yes, 0 = no)

Is the background of the page solid or non-solid? (1 = solid, 2 = non-solid)

Is the contrast between text and background high or low? (1 = high, 2 = low)

How much clutter is there on the page? (1 = clutter, 2 = no clutter)

Is the density on the page high or low? (1 = high, 2 = low)

Is there a clear focal point presented (e.g. a large image or a large title) (1 = yes, 0 = no)

1c. Readability

What is the font size? (1 = normal, 2 = small, 3 = large)

Is the font playful? (1 = yes, 0 = no)

Are there long or short texts on the page? (1 = short texts, 2 = long texts)

Are subheadings presented in the text on the page? (1 = yes, 0 = no)

1d. Type of Media

How many images are presented on the page? (0 = no images, 1 = low number of images, 2 = average number of images, 3 = high number of images)

Are there animations or moving images or objects presented on the website? (1 = yes, 0 = no)

Are there videos presented on the website? (1 = yes, 0 = no)

Do they make use of sounds on the website? (1 = yes, 0 = no)

· Do these sounds start immediately when the user enters a page? (1 = yes, 0 = no)

· Do these sounds start when the user hovers or clicks on a link? (1 = yes, 0 = no)

· Can the user control the sounds by turning them on or off? (1 = yes, 0 = no)

· Are the sounds functional or non-functional? (1 = functional, 2 = non-functional)

1e. Layout

Is the layout of the web pages consistent throughout the website? (1 = yes, 0 = no)

2. Navigation

2a. Type of navigation tools

Are there menu structures presented with main and subcategories (1 = yes, 0 = no)

Is there a search engine available? (1 = yes, 0 = no)

Does the logo also work as a home button? (1 = yes, 0 = no)

Is there a separate home button presented? (1 = yes, 0 = no)

2b. Format of the menu

Is there a menu presented in a row of textual or image labels? (1 = yes, 0 = no)

· Are the menu items presented by…

o textual labels? (1 = yes, 0 = no)

by textual labels with images? (1 = yes, 0 = no)

o only by images? (1 = yes, 0 = no)

· Are the images used for the menu items meaningful? (1 = meaningful, 2 = not meaningful)

Is the menu presented within an Image map (i.e. a large image with clickable objects that represent the main categories)? (1 = yes, 0 = no)

2c. Appearance of new web page

Does a new page appear…

· in the same window? (1 = yes, 0 = no)

· in a new window? (1 = yes, 0 = no)

· to a bookmark on the same page? (1 = yes, 0 = no)

2d. Marking of hyperlinks

Are hyperlinks recognizable as clickable? (1 = yes, 0 = no)

Are hyperlinks marked as clickable when the user hovers over them with the mouse (e.g. by changing size or color of the object)? (1 = yes, 0 = no)

2e. Use of orientation cues

Is the user's location within the website marked in the main menu? (1 = yes, 0 = no)

Is the user's location within the website marked by a back ground color used for that particular section of the website? (1 = yes, 0 = no)

Is a navigation path presented at the top of the page (i.e. a crumbs trail)? (1 = yes, 0 = no)

2f. Search engine

Does the search engine find results…

· on the website itself? (1 = yes, 0 = no)

· on other websites? (1 = yes, 0 = no)

· on the entire 'mother website' (for example in case of a 'kids corner')? (1 = yes, 0 = no)

2g. Characteristics of the search results

What are the characteristics of the presented search results?

· Does it only consist of a title of the result page? (1 = yes, 0 = no)

· Does it present a summary of the result page? (1 = yes, 0 = no)

· Does it present the theme to which the result page belongs? (1 = yes, 0 = no)

· Does it present a description of the result page in one sentence? (1 = yes, 0 = no)

· Does it present a snippet of the text from the result page? (1 = yes, 0 = no)

· Are relevant keywords for the search query highlighted? (1 = yes, 0 = no)

· Is a thumbnail presented of the result page? (1 = yes, 0 = no)

· How many results are presented on the first page with search results?

· Is there an opportunity to go to more search results on next search results pages? (1 = yes, 0 = no)

· Does the search engine provide query suggestions when the user types in a search query? (1 = yes, 0 = no)

· Does the search engine provide a 'Did you mean' tool when the spelling is incorrect? (1 = yes, 0 = no)

· Does the search engine provide search help when no results are found? (1 = yes, 0 = no)

· Does the search engine accept keywords? (1 = yes, 0 = no)

· Does the search engine accept natural language (e.g. an interrogative sentence)? (1 = yes, 0 = no)

2h. Menu structure

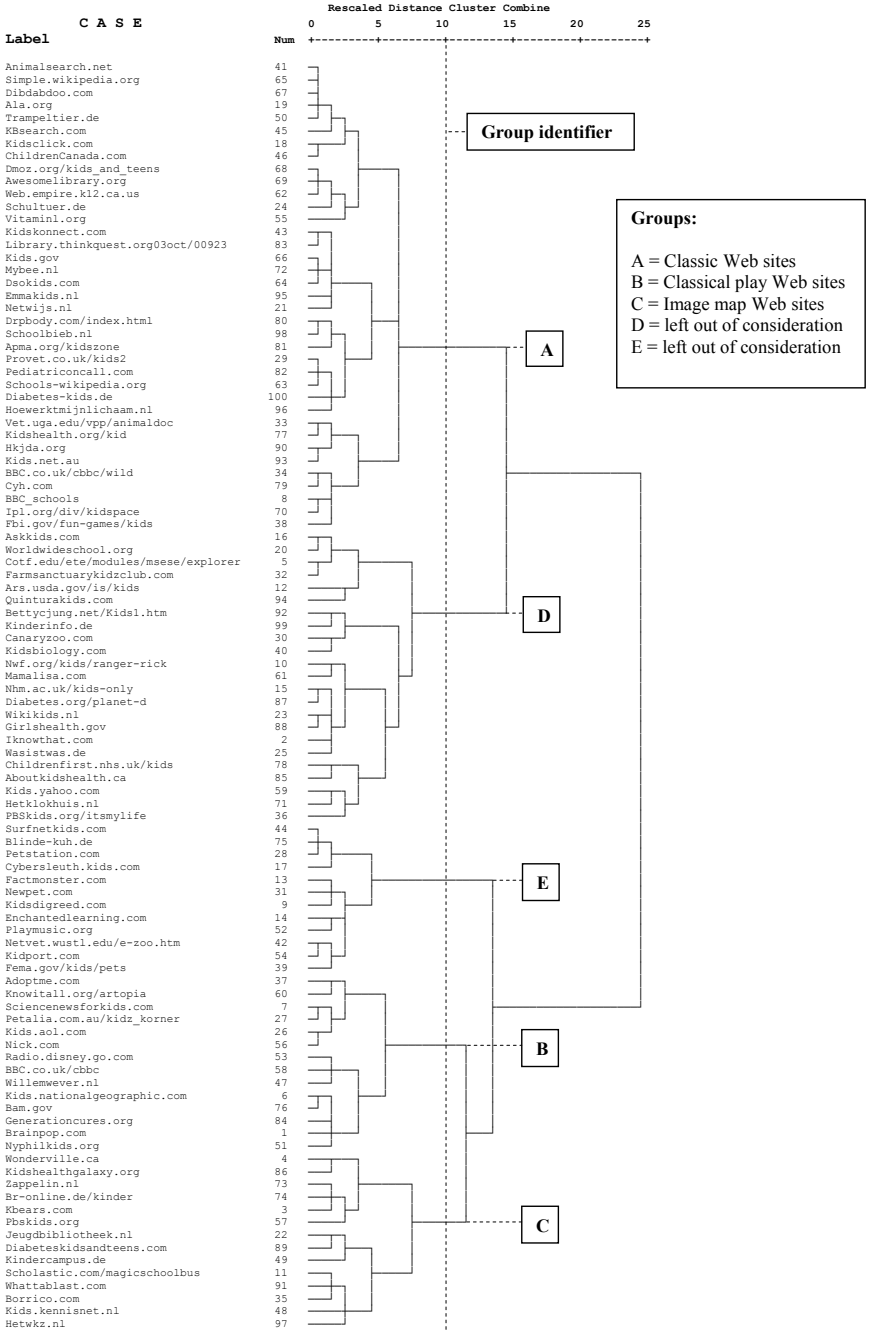
Is the menu structure consistent throughout the website? (1 = yes, 0 = no)

Is the menu structure user friendly? (1 = yes, 0 = no)

Are the main categories provided concrete or ambiguous? (1 = concrete, 2 = ambiguous)

Appendix B: Dendrogram as output of the hierarchical cluster analysis (chapter 3)

Dendrogram using Ward Method



Appendix C: Search tasks per website (chapter 4)

Themes	Classic website	Classical play website	Image map website
Animals	**Kangaroo:** Kangaroos mostly live in Australia. Baby kangaroos grow up inside of the mother's pouch. How does the baby kangaroo stay put inside of the pouch while the mummy jumps around all the time? Can you find an answer to this question on the website schoolbieb.nl?	**Giraffe:** Giraffes have a very long neck, but what is actually the reason for having such a long neck? Can you find an answer to this question on the website Willemwever.nl?	**Sharks:** There are many different types of sharks. Luckily, most of them are not dangerous for human beings. However, because of their enormous number of sharp teeth, they look very dangerous to us. When sharks loose a teeth, they automatically get a new one. Can you find out on the website kids.kennisnet.nl how this works?
Painters and composers	**Beethoven:** Beethoven was a famous composer, who wrote lovely music. Even when he was entirely deaf, he could still write music. Can you find the name of one of his famous plays on the website schoolbieb.nl?	**Rembrandt:** Rembrandt was a famous painter. One of his most famous paintings is named 'de Nachtwacht'. Can you find out on the website willemwever.nl why he painted this painting in the first place?	**Vincent van Gogh:** The paintings of Vincent van Gogh became famous a long time after he passed away. Can you find out on the website kids.kennisnet.nl how the painting style is named that Van Gogh applied in his later works?
Sports	**Basketball:** Basketball is a very popular sports game in the United States of America. The scoring in this game is a bit odd, because when you score in basketball you can get one, two or three points. Can you find out on the website schoolbieb.nl when you score three points in basketball?	**Gymnastics:** Gymnasts often train many hours a day. They often look smaller than other sports men and women. Can you find out on the website willemwever.nl whether this is a coincidence or whether this is caused by their hard work as a gymnast?	**Hockey:** Hockey is a popular sports game in the Netherlands. The game is played by eleven players. The eleventh player is the keeper. Can you find out on the website kids.kennisnet.nl why the keeper wears so much stuff during the game and what kind of stuff he wears?
Health	**Brackets:** Many children get brackets to straighten their teeth. Can you find out on the website schoolbieb.nl how long it usually takes before a bracket is removed when the teeth are straightened?	**Nausea:** Can you find out on the website willemwever.nl why you can get sick by reading something while sitting in a moving car?	**Head lice:** It is very annoying when you have lice. It does not mean that you are filthy when you have head lice, because lice like clean heads. Can you find out on the website kids.kennisnet.nl why it itches terribly when you have lice?

Appendix D: Task descriptions translated from the original Dutch versions (chapter 5)

Task 1
You've seen a chameleon in the zoo and you saw that he was moving very slowly.
Now you wonder how a chameleon captures its prey, because he seems far too lazy for that.
Can you find out how a chameleon catches its prey at Junior Winkler Prins online?
Optimum navigation path:
Correct main category: Nature
Correct subcategory: Reptiles and amphibians
Correct content page: Chameleon

Task 2
You would like to become an astronaut and explore space in search of alien planets, like in the movies. Can you find out who was actually the first living creature that traveled through space at Junior Winkler Prins online ?
Optimum navigation path:
Correct main category: Science and technology
Correct sub category: Space and Space travel
Correct content pages: Space / Astronaut

Task 3
You spend your holiday in Friesland with your parents and your father told you that he participated in the famous Dutch skating tour named 'Elfstedentocht' once.
He also told you that he met the Dutch crown prince Willem-Alexander during the tour.
Can you find out under what name Willem-Alexander participated in the scating tour that day?

Optimum navigation path:
Correct main category: Sports and exercise
Correct subcategory: Stadiums & Tournaments
Correct content page: "Elfstedentocht"

Task 4
It is often said that Christopher Columbus discovered America. Some say that this is not true. They say that another nation discovered America 500 years before Columbus did.
Can you find out what people discovered America before Columbus did at Junior Winkler Prins online?
Optimum navigation path:
Correct main category: History
Correct subcategories: Time of cities and states / Time of discoverers and reformers / Time of monks and knights
Correct content pages: Vikings / Discoveries

Task 5
Isaac Newton invented a device to see the stars better. Can you find out what the name of that device is at Junior Winkler Prins online?
Optimum navigation path:
Correct main category: Famous people / Science and technology
Correct subcategories: Inventors and scientists / Space and space travel / Measuring, weighing or counting / How it works

Correct content pages: Isaac Newton / Star / Telescope

Appendix E: Bipolar picture anchors for pragmatic and hedonic quality from the AttracDiff 2 questionnaire (chapter 5)

Pragmatic quality (PQ)

Figure 1. Technical – human

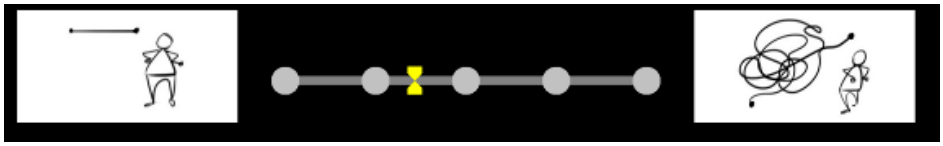


Figure 2. Simple - complicated

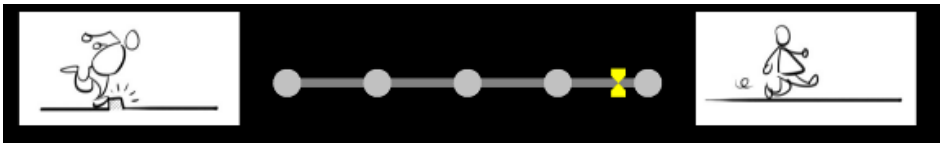


Figure 3. Impractical – practical

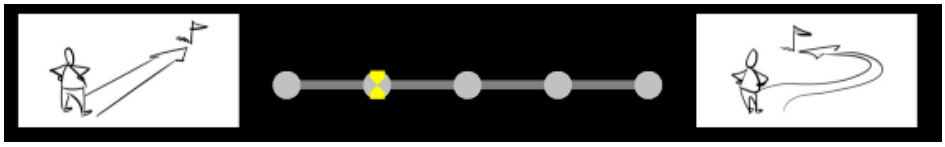


Figure 4. Direct - cumbersome

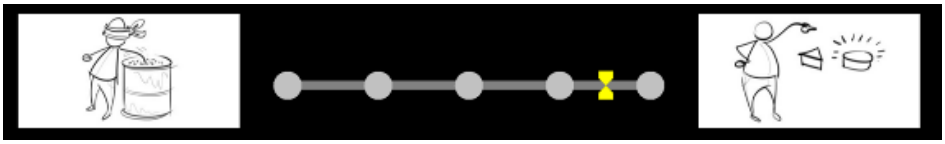


Figure 5. Unpredictable – predictable*

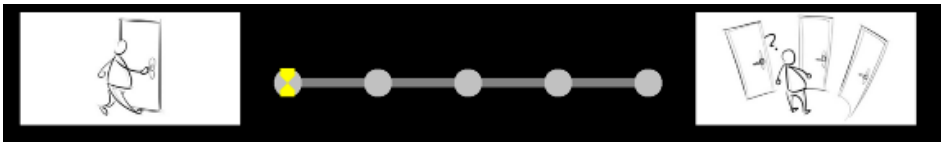


Figure 6. Clear - confusing

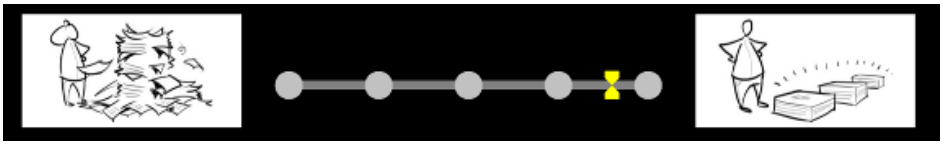


Figure 7. Unruly – manageable

Hedonic quality - identification (HQI)

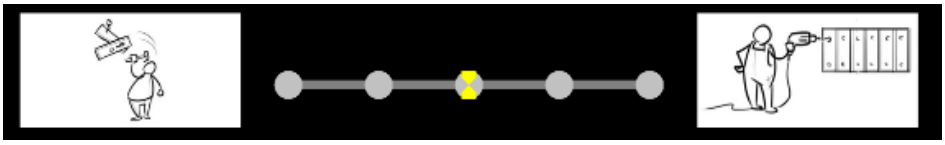


Figure 8. Amateurish – professional

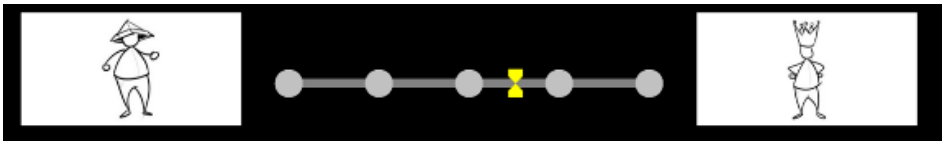


Figure 9. Cheap – valuable

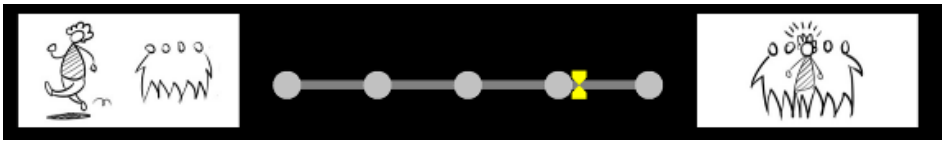


Figure 10. Non-inclusive – inclusive*

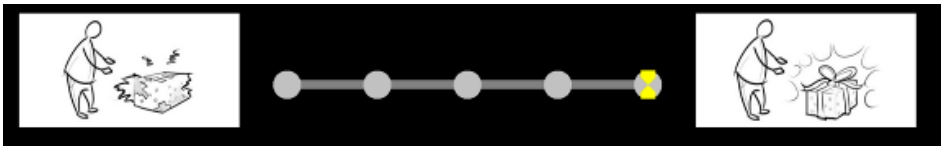


Figure 11. Unpresentable – presentable

Hedonic quality – stimulation (HQS)

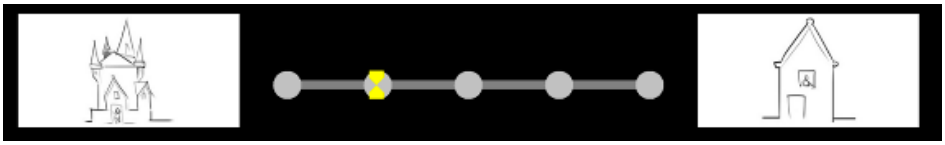


Figure 12. Creative - standard*

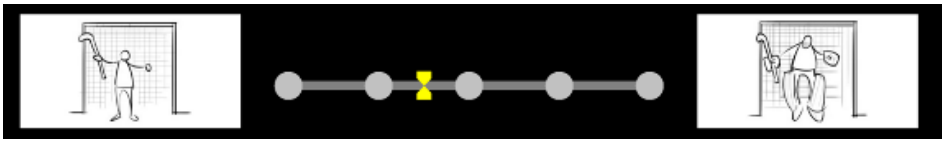


Figure 13. Courageous - cautious*

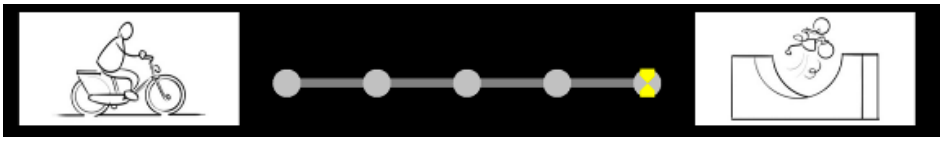


Figure 14. Lame – exciting

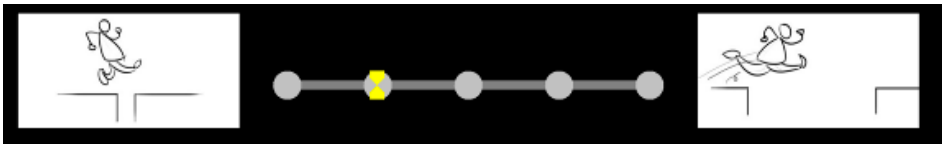


Figure 15. Easy – challenging

** Picture anchors that were left out of the final experiment based on the results of the pilot study*

Appendix F. Mean scores for the variables to measure perceptions of pragmatic and hedonic quality (5-point semantic differential scales) for the different versions of the website (SD in parentheses).

	Pragmatic Quality						*Hedonic Quality*				
Website version	*Technical-human*	*Simple-Complcated*	*Impractical-practical*	*Direct-cumbesome*	*Clear- con-fusing*	*Unruly-manageable*	*Amateurish-profesional*	*Cheap-valuable*	*Unpresent-able presentable*	*Lame-exciting*	*Easy-chal-lenging*
Classic (N = 51)	3.27 (1.21)	2.29 (1.20)	3.86 (1.01)	2.58 (1.13)	2.75 (1.42)	4.25 (.95)	4.24 (.71)	3.77 (.85)	4.14 (.87)	3.39 (1.13)	2.50 (1.17)
Classical play (N = 52)	3.39 (1.14)	2.35 (1.02)	3.78 (.89)	2.49 (.98)	2.58 (.97)	4.10 (.79)	3.88 (.90)	3.66 (.79)	4.17 (.74)	3.61 (.96)	2.81 (1.01)
Image map (N = 55)	3.28 (.86)	2.90 (1.11)	3.53 (.95)	2.97 (.93)	3.26 (1.09)	3.90 (.91)	3.90 (.78)	3.52 (.98)	3.87 (.78)	3.60 (1.00)	3.02 (1.01)

*There is a significant effect of the type of design on the affective responses.

Mean scores for the variables to measure emotional valence (1= very negative feeling; 5 = very positive feeling), emotional arousal (1 = not emotionally aroused at all; 5 = very emotionally aroused) and the product evaluations for beauty, goodness and fun (report mark from 1-10) for the different versions of the website (SD in parentheses).

Website version	*Valence*	*Arousal*	*Goodness*	*Fun*	*Beauty*
Classic (N = 51)	4.12 (.77)	2.55 (1.30)	8.66 (1.31)	8.29 (1.56)	8.37 (1.42)
Classical play (N = 52)	4.38 (.66)*	2.73 (1.29)	8.71 (.84)	8.50 (1.19)	8.86 (1.05)
Image map (N = 55)	4.05 (.68)	2.80 (1.21)	8.22 (1.04)*	8.13 (1.10)	8.42 (1.03)

*There is a significant effect of the type of design on the affective responses.

Publications

Jochmann-Mannak, H., Lentz, L., Huibers, T., & Sanders, T. (2014). How interface design and search strategy influence children's search performance and evaluation. In D. Yannacopoulos, P. Manolitzas, N. Matsatsinis, & E. Grigoroudis (Eds.), *Evaluating Websites and Web Services: Interdisciplinary Perspectives on User Satisfaction* (pp. 241-287). Hershey, PA: IGI Global. (Chapter 5)

Jochmann-Mannak, H., Lentz, L., Huibers, T., & Sanders, T. (2012). Three types of children's informational Websites: An inventory of design conventions. *Technical Communication*, *59*(4), 302-323. (Chapter 3)

Jochmann-Mannak, H., Huibers, T., Lentz, L., & Sanders, T. (2010). Children searching information on the Internet: Performance on children's interfaces compared to Google. Paper presented at the 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval at the workshop 'Towards Accessible Search Systems', Geneva, Switzerland. (Chapter 4)

Jochmann-Mannak, H., Huibers, T. & Sanders, T. (2009). Children's information retrieval: how to support children in effective information seeking. Paper presented at the 9th Dutch-Belgian Information Retrieval Workshop, Enschede, The Netherlands, CTIT Workshop Proceedings, ISSN 0929-0672. (Chapter 2)

Jochmann-Mannak, H., Huibers, T. & Sanders, T. (2008). Children's information retrieval: beyond examining search strategies and interfaces. Paper presented at the 2nd BCS IRSG Symposium: Future Directions in Information Access, London, UK. (Chapter 2)

Summary

Every day, more children use digital media to search for information. While most children aged 8-12 make use of Google to search for information, research shows that children experience all kinds of problems using search interfaces such as Google. One of the reasons is that these informational search interfaces were not designed for children.

The research reported in this dissertation is about the design of informational interfaces for children. We explored how children interact with these interfaces. And we tested the effects of specific design choices on children's search performance and their attitude towards different interface design types. The goal of the different studies reported in this dissertation was to learn more about children's interaction with different interface design types and to learn how to design interfaces that 'work' for children and that children 'like'. We expected that both usability and the fun factor of an informational website are important for children.

In this research we made use of a mixed methods design in which both quantitative and qualitative research methods were mixed. Three iterative steps were taken following one another in time in which one stage influenced the decision about the next stage. The first step was a corpus study to identify design conventions of children's search interfaces. The second step was an explorative study to explore children's search behaviour on existing search interfaces that are representative for the children's search interfaces, based on the results of the corpus study. The third step was an experiment to test children's search performance on and affective responses towards three types of search interfaces for children. The interfaces for this experiment were developed based on the results of the corpus study. The research directions suggested by the results of the explorative study were empirically tested in this experiment.

The corpus study

Prior research on web design conventions has an almost exclusive focus on web design for adults. There is far less knowledge about web design for children. In the corpus study, we presented an overview of the current design conventions for children's informational websites. We analyzed a large corpus of 100 children's international, informational websites from four different domains (i.e. science, pets, arts and health). The instrument used for analyzing the websites included categories on visual design, navigation and information architecture. The design conventions identified in this study showed that designers of children's informational websites often follow general web design guidelines. An overview is presented of the current design conventions for children's informational websites.

A closer look at the data revealed three categories of informational websites especially designed for children, diverging from a classical to a playful design approach:

1) a classical interface design with a classical interaction style and without playful graphics, named 'Classic web design'
2) an interface design with playful graphics, but a classical interaction style, named 'Classical play web design'
3) a playful interface design with playful graphics and a playful interaction style, named 'Image map web design'.

The effects of playful design on children's search performance and affective responses formed an important topic in the next stages of the research reported in this dissertation.

The explorative study

To explore children's search behaviour on and attitude towards different types of website design for children, a study was conducted with 32 children in the age range of 8 to 12 years old. We explored the effects of different design characteristics on performance, evaluation and experienced emotions.

The children conducted search tasks on five different websites: four websites for children and Google. Video and audio recordings were made of their screen activity and their emotional expressions. We also recorded children's eye-movements over the screen during task performance using an eye-tracking device. Our main interest went to three of the websites which represented the three types of children's web design identified in the corpus study.

The Image map website (playful in both visual and navigation design) scored significantly lower for all pragmatic variables (time, clicks and success) than the other website types. Furthermore, the Image map website evoked more emotional expressions than the other two websites, both positive and negative. This was primarily caused by aesthetic reasons (for example, beauty and fun).

These explorative results suggested that playful navigation lowers usability of a website but playful visual design seemed to have a positive effect on perception of hedonic quality. Furthermore, this study provided an evaluation of different methods that can be used in research on children's search performance and perceptions of pragmatic (i.e. whether children find it 'easy to work' with) and hedonic quality (i.e. whether children 'like' it).

The experiment

Finally, we conducted an experiment with 158 children, aged 10 to 12, in which search performance and attitudes towards an informational website were investigated to provide evidence for the effects of interface design suggested by the results of the explorative study. The same website (i.e. Junior Winkler Prins online encyclopedia) was designed in three different types of interface design varying in playfulness of visual design and in playfulness of navigation structure. We also manipulated the used search method (i.e. browsing or searching) by providing a website version to the children with or without a search engine.

We found that the type of interface design did not have an effect on the children's search performance (in terms of search success, amount of time and number of clicks) but it did influence the children's feelings of emotional valence and their evaluation of 'goodness'. The children felt most positive about the Classical play website with a classical navigation structure and playful aesthetics. They found the playful Image map website least good of the three website versions. Furthermore, the search performance of the children was much more effective and efficient when using the search engine than when browsing the menu. The research also showed how difficult it is to measure affective responses towards digital interfaces with children.

Conclusion

The goal of the research was to learn more about interfaces that 'work' for children and interfaces that children 'like'. Our research results show that if developers want to support children in effective and efficient information search, they should be careful with the use of playful navigation design (such as the Image map). Children appreciate playful design, preferably when it concerns visual design elements (such as the Classical play website). Playful design elements may not be at the expense of usability for children, because ease of use is more important for children when searching for information, than having a playful experience. Furthermore, providing a search engine is an absolute necessity to optimally support children in effective and efficient search performance.

Samenvatting in het Nederlands

Steeds meer kinderen maken gebruik van digitale zoekdiensten om naar informatie te zoeken. De meeste kinderen tussen de acht en twaalf jaar maken gebruik van Google, maar uit verschillende onderzoeken blijkt dat kinderen veel problemen ervaren met het gebruik van zoekdiensten zoals Google. De belangrijkste oorzaak van deze problemen is dat dergelijke informatieve websites niet zijn gemaakt voor kinderen.

Het onderzoek dat beschreven staat in dit proefschrift gaat over het ontwerp van informatieve websites voor kinderen. Wij hebben onderzocht hoe kinderen omgaan met verschillende typen websites, waarbij wij wilden weten wat voor soort websites kinderen gemakkelijk konden gebruiken. Daarnaast hadden we de verwachting dat de mate waarin kinderen een website leuk vinden een belangrijke factor is voor kinderen. Daarom wilden we ook weten wat voor soort websites kinderen wel en niet leuk vinden om mee te werken.

Wij hebben in het onderzoek gebruik gemaakt van zowel kwantitatieve als kwalitatieve onderzoeksmethoden. In totaal hebben we drie opeenvolgende stappen doorlopen, waarbij de resultaten van de ene stap steeds werden gebruikt werd om de volgende onderzoeksstap te bepalen. De eerste stap was een corpusonderzoek, waarbij we de conventies voor het ontwerp van informatieve kinderwebsites in kaart hebben gebracht. De volgende stap was een verkennend onderzoek, waarbij we het zoekgedrag van kinderen hebben bestudeerd op bestaande websites. Op basis van de resultaten uit het corpusonderzoek en het verkennende onderzoek hebben we een experiment uitgevoerd. Hierbij hebben we het verschil tussen drie verschillende typen kinderwebsites onderzocht; enerzijds in termen van de effectiviteit en efficiëntie van de taakuitvoering van kinderen en anderzijds in termen van hun waardering van de websites en hun emotionele reacties.

Het corpusonderzoek

Onderzoek naar conventies voor webontwerp gaat in bijna alle gevallen over websites voor volwassenen. Wij weten veel minder over de conventies die worden gebruikt bij websites voor kinderen. Daarom hebben wij in een corpusonderzoek een overzicht gegeven van de conventies voor het ontwerp van kinderwebsites. We hebben een corpus van honderd internationale, informatieve websites voor kinderen met betrekking tot vier thema's geanalyseerd: wetenschap, huisdieren, kunst en gezondheid. In het gehanteerde analyse-instrument waren onder andere de categorieën visueel ontwerp, navigatie en informatiearchitectuur opgenomen. Het bleek dat de conventies die worden

gebruikt voor het ontwerpen van kinderwebsites zijn gebaseerd op de algemene conventies voor het ontwerpen van websites.

Uit verder onderzoek van de data bleek dat er drie verschillende typen websites voor kinderen konden worden geïdentificeerd, die uiteenliepen van een klassieke tot een speelse manier om websites voor kinderen vorm te geven. Deze drie typen kunnen we als volgt beschrijven:

1) Klassiek webontwerp zonder visueel speelse kenmerken (zoals plaatjes en animaties) en met een klassieke navigatiestructuur.
2) Speels-klassiek webontwerp met visueel speelse kenmerken (zoals plaatjes en animaties), maar een klassieke navigatiestructuur.
3) 'Image map' webontwerp waarbij zowel de visuele kenmerken als de navigatiestructuur speels zijn opgezet.

Op basis van dit onderzoek waren wij benieuwd of er verschillen zouden zijn in de effectiviteit en efficiëntie waarmee kinderen informatie kunnen zoeken met deze verschillende websitetypen. Ook vroegen wij ons af of er een verschil zou zijn in de waardering van kinderen van deze verschillende typen webdesign. Dit vormde de aanleiding voor de volgende onderzoeksstappen.

Het verkennende onderzoek

Nadat we in kaart hadden gebracht wat voor soort informatieve websites er allemaal zijn voor kinderen, wilden we verkennen hoe kinderen zoal te werk gaan op websites die speciaal voor hen zijn ontworpen. Aan dit verkennende onderzoek werkten 32 kinderen mee in de leeftijd van acht tot twaalf jaar. Hierbij keken we naar de verschillen in taakuitvoering, waardering en emoties van de kinderen tussen de verschillende websites. Ieder kind voerde vijf taken uit op vijf verschillende websites. Vier van deze websites zijn speciaal ontworpen voor kinderen (de klassieke website schoolbieb.nl, de speels-klassieke website willemwever.nl, de Image map website kids.kennisnet.nl en wikikids.nl). De vijfde website was Google. Tijdens de taakuitvoering maakten we video- en audio-opnamen om de taakuitvoering nauwkeurig te kunnen analyseren. Ook maakten we opnamen van de oogbewegingen over het scherm van de kinderen tijdens de taakuitvoering.

De kinderen bleken het minst effectief en efficiënt te kunnen werken met de Image map website. Deze Image map website veroorzaakte de meeste emotionele reacties, waarbij ongeveer evenveel positieve als negatieve emotionele reacties werden geuit. Deze reacties bleken met name betrekking te hebben op hoe mooi of leuk de kinderen de website vonden.

De resultaten uit het verkennende onderzoek gaven ons het idee dat door het toevoegen van speelse elementen aan de navigatiestructuur de gebruiksvriendelijkheid van websites vermindert. Aan de andere kant leken visueel speelse elementen (zoals plaatjes en animaties) juist een positief effect te hebben op de mate waarin kinderen de websites mooi of leuk vonden. Om deze mogelijke effecten van speels webontwerp daadwerkelijk aan te kunnen tonen, hebben we vervolgens een experiment opgezet.

Het experiment

Wij hebben een experiment uitgevoerd waaraan 158 kinderen in de leeftijd van tien tot twaalf jaar hebben meegewerkt. Voor dit experiment hebben we drie verschillende versies van de Junior Winkler Prins online encyclopedie gemaakt. Deze drie versies representeerden de drie verschillende typen kinderwebsites, zoals geïdentificeerd in het corpusonderzoek. We manipuleerden ook de zoekmethode die de kinderen moesten gebruiken door een versie met of zonder zoekmachine aan te bieden.

Er bleek geen verschil te zijn tussen de drie verschillende typen kinderwebsites in termen van taaksucces en de benodigde hoeveelheid tijd en het aantal kliks. Voor de effectiviteit en efficiëntie van de taakuitvoering maakte het dus niet uit of de kinderen gebruikmaakten van een klassieke, speels-klassieke of Image map versie van de website. Er was wel verschil in het gevoel dat de kinderen bij de verschillende websites hadden en in de mate waarin ze de websites 'goed' vonden. De kinderen waren namelijk het meest positief gestemd over de speels-klassieke versie met een klassieke navigatiestructuur en speelse visuele elementen (zoals plaatjes). Van de drie versies vonden de kinderen de Image map versie het minst goed. Het onderzoek toonde ook aan dat het gebruik van een zoekmachine om naar informatie te zoeken effectiever en efficiënter werkte voor de kinderen dan het zoeken naar informatie via het hoofdmenu. Ten slotte lieten we met dit onderzoek zien hoe moeilijk het is om gevoelens over en waardering van een website te meten met kinderen.

Conclusie

Dit onderzoek had als doel erachter te komen met wat voor soort websites kinderen gemakkelijk uit de voeten kunnen en wat voor soort informatieve websites kinderen wel en niet leuk vinden om mee te werken. Ons onderzoek toont aan dat ontwikkelaars van kinderwebsites erg voorzichtig moeten zijn met het gebruik van speelse navigatie-elementen (zoals een Image map). Kinderen waarderen speelse elementen in websites, vooral wanneer het om visueel speelse elementen gaat zoals plaatjes en animaties.

Kinderen waarderen het zeker niet wanneer het toevoegen van speelse elementen ten koste gaat van de gebruiksvriendelijkheid van een website. Gebruiksvriendelijkheid van een website is belangrijker voor kinderen wanneer zij op zoek zijn naar informatie dan de mate waarin zij een website mooi of leuk vinden vanwege speelse elementen. De belangrijkste conclusie uit ons onderzoek: om kinderen optimaal te ondersteunen in effectief en efficiënt zoekgedrag is het aanbieden van een zoekmachine op een kinderwebsite een absolute noodzaak.

Dankwoord

Het zit erop! Als mensen ernaar vragen kan ik eindelijk vertellen dat mijn proefschrift klaar is. Dan kan ik ze ook met veel trots dit mooie boek aanbieden als resultaat van mijn onderzoek. Als ik het helemaal alleen had moeten doen, had dit boek er zeker niet geweest. Daarom wil ik van deze gelegenheid gebruik maken om al die mensen te bedanken die een bijdrage hebben geleverd aan dit onderzoek, op welke wijze dan ook. Als je door dit boek bladert, zul je wel merken dat ik nogal een kletskous ben. Ik ben niet van het bondige schrijven. Zo is dit dankwoord ook niet kort en bondig.

Promotoren, bedankt!

Ik wil graag beginnen met het bedanken van mijn drie promotoren, Theo Huibers, Leo Lentz en Ted Sanders. Ik denk niet dat er veel doctors kunnen zeggen dat ze door drie professoren zijn begeleid. Dankzij jullie hulp ligt hier nu een prachtig boek.

Theo, dankjewel dat je me de kans hebt gegeven om dit promotieonderzoek uit te voeren. Dankjewel voor de vele gesprekken en mailwisselingen, je visie, je kritische noot en je altijd terugkerende les dat het glas niet half leeg is, maar half vol! Heel bijzonder hoeveel tijd je hebt vrijgemaakt voor mijn begeleiding, terwijl je zo'n drukke agenda hebt met grote klanten en raden van bestuur. Maar toch maak je net zoveel tijd voor een 'Twentse' aio. Ook de leuke gesprekken over de kindertjes, de hond of de vakanties zal ik nooit vergeten. Of wat dacht je van de bokmuizen of de ritjes naar het station in de Audi.

Leo, na een jaar werd het mij duidelijk dat een promotieonderzoek voor mij niet was weggelegd zonder dagelijks begeleider. Ik ben je eeuwig dankbaar dat je mij te hulp schoot om mijn onderzoek concreet vorm te geven. Hoe doe jij dat toch? Als ik even helemaal geen idee had hoe ik verder moest, wist jij altijd weer binnen een mum van tijd de boel helder op een rijtje te krijgen. Na een gesprek met jou wist ik ineens weer precies wat ik aan het doen was en wat mijn vervolgstappen werden. Sorry dat dit jou wel regelmatig een pakje zakdoeken en een geschild appeltje kostte, maar ook daarvoor ben ik je heel dankbaar. Ten slotte, dank je voor al je hulp bij het schrijven en reviseren van mijn artikelen, waarbij je soms wel tien versies van hetzelfde stuk hebt moeten lezen.

Ted, dankjewel dat je als hoofd Taalbeheersing enthousiast werd over dit onderzoek, ondanks dat het niet het soort onderzoek was dat je gewend was te begeleiden. Dank je dat jij mij de mogelijkheid hebt gegeven mijn onderzoek als 'visiting PhD' aan de Trans in Utrecht uit te mogen voeren. Het feit dat jij in de gesprekken altijd de rode draad wist te benoemen - en vooral ook te bewaken - was van enorme waarde. Zeker bij het schetsen van het grotere kader in de laatste fase van het schrijven van mijn proefschrift was jouw inbreng van goud.

Aan allen die mijn onderzoek mede mogelijk maakten, bedankt!

Natuurlijk viel er weinig te onderzoeken als er geen scholen en kinderen op scholen bereid waren mee te werken aan mijn onderzoek. Daarom wil ik alle docenten en kinderen bedanken die aan mijn onderzoek hebben meegewerkt.

Zonder jouw hulp, Huub van den Bergh, had ik bijna geconcludeerd dat er een significant verschil was in taakuitvoering tussen de verschillende designtypen in mijn experiment. Gelukkig schoot jij te hulp om met mij een multilevel analyse te doen waaruit bleek dat ik dit helemaal niet kon concluderen uit de verzamelde data. Dankjewel voor al je hulp bij de statistische analyses in mijn proefschrift.

En hoe gaaf was het om in mijn experiment ook fysiologische data bij de kinderen te kunnen meten. Dankjewel Matthijs Noordzij, dat jij mij de mogelijkheid bood gebruik te maken van de Q-sensors, voor je hulp bij het opzetten van het afnameprotocol en voor je hulp bij het analyseren van alle Q-data. Nu bleek het uitvoeren van een schoolse taak achter een computer erg weinig opwindend voor de kinderen, maar het was zeker het proberen waard!

Iris Mulders, dankjewel voor de mogelijkheid om met de Tobii eye-tracker aan de slag kunnen bij mijn eerste onderzoek en bedankt voor al je hulp bij het 'leren kennen' van de Tobii eye-tracker.

Ik wil de medewerkers van Uitgeverij Unieboek | Het Spectrum bv die zich bezighouden met de Junior Winkler Prins online hartelijk danken dat ik deze website mocht gebruiken in mijn experiment en voor het ontwikkelen van de webomgeving voor mijn experiment. Ook wil ik Mark en Michelle van internetbureau Vierkante Meter bedanken voor het ontwerpen van de klassieke en speels-klassieke versie van de Junior Winkler Prins online voor mijn experiment.

Verder wil ik alle bachelor- en masterstudenten bedanken die hebben bijgedragen aan het analyseren van de enorme hoeveelheid kwantitatieve en kwalitatieve data die ik heb verzameld in de verschillende studies.

Aan allen die mijn leven als aio nog leuker maakten, bedankt!

Ik ga nu ongeveer in chronologische volgorde de mensen langs die mijn leven als aio nog leuker hebben gemaakt gedurende de afgelopen jaren. Nina, dankjewel dat je me hebt geattendeerd op de vacature van dit onderzoek. Zonder jou had ik geen weet gehad van het bestaan van dit leuke project. Ook heb jij mij na mijn studie geïntroduceerd in de wereld van onderzoek doen en aio zijn. Dankjewel voor de gezellige tijd samen op de zolderkamer naast Huub en voor je betrokkenheid de afgelopen jaren. Mijn onderzoek begon bij de dames van de afdeling Educatie van de Bibliotheek van Rotterdam. Dank jullie wel voor jullie gastvrijheid. Ik mocht een jaar bij jullie werken, aangezien Twente wel erg ver is voor een Rotterdammer. Ondanks dat het niet handig

bleek om als startende aio niet tussen andere onderzoekers te zitten, heb ik een erg leuke en gezellige tijd bij jullie doorgebracht.

Na een jaar in de Rotterdamse bibliotheek kon ik gelukkig aan de slag tussen andere onderzoekers bij de afdeling Taalbeheersing in Utrecht. Aangezien ik voor mijn gevoel al een heel jaar had 'weggegooid', zag ik het eigenlijk helemaal niet meer zitten om nog verder te gaan. Sanne, ik ben je eeuwig dankbaar dat jij mij er toen doorheen hebt gesleept. Je hebt me op weten te beuren en Leo ingeschakeld voor een reparatieslag. Ik heb ontzettend veel van jou geleerd als onderzoeker, je was een hele fijne en lieve collega waar ik altijd bij terecht kon, je sprak me streng toe als dat even nodig was en bovenal heb je me leren relativeren! Sanne, ik ben daarom blij en dankbaar dat jij mijn paranimf bent.

Het begin van mijn tijd in Utrecht zat een deel van de afdeling Taalbeheersing aan het Janskerkhof. Het pand was niet erg fijn of gezellig, maar Hanneke, wat vond ik ons piepkleine kamertje waar ik samen met jou mocht zitten gezellig! Dankjewel voor de leuke tijd die we samen als collega's hebben doorgebracht. We hebben hard gewerkt, maar er was ook regelmatig tijd voor een welverdiend kopje thee met een gezellig gesprek. Ik heb je wel gemist hoor, toen je na twee jaar bij ons stopte om met Molenaar Communicatie te starten.

In de eerste paar jaar van mijn onderzoek, bracht ik regelmatig één of twee dagen door bij mijn afdeling HMI aan de UT in Twente. Wim, Ivo, Thijs, Bart en Herwin, bedankt dat jullie mij ook een beetje 'thuis' lieten voelen in het verre Enschede. Ronald, Mark en Frans, bedankt voor de goede gesprekken tijdens mijn bezoeken aan de UT in de afgelopen tijd toen al mijn Twentse 'kamergenoten' inmiddels al waren vertrokken. Charlotte, dankjewel voor je hulp op afstand bij alle administratieve zaken die als Twentse aio op mijn pad kwamen. En - last but not least - Lynn, heel veel dank voor het 'nakijken' van het Engels in mijn proefschrift. Zal ik ooit het verschil tussen 'that' en 'which' goed gaan begrijpen?

Mijn laatste verhuizing binnen de UU was misschien niet het beste voor mijn proefschrift. Véél te gezellig! Ik denk dat het niet voor niets is dat ik vanaf die tijd veel meer thuis ben gaan werken. Maar oh wat fijn om zo nu en dan in Utrecht te zijn om te sparren met mijn kamergenoten, of om gewoon collega's om me heen te hebben. Naomi, wat hebben we een lol gehad. Jij straalt zoveel warmte uit. Maar ook voor zoveel onderzoeksgerelateerde vragen kon ik bij jou terecht. Marion, dankjewel dat ik jouw paranimf mocht zijn. Bijzonder om dat een keer mee te mogen maken. Dankjewel voor je nuchterheid en betrokkenheid.

Gerdineke, Rogier en Monica, dank jullie wel voor de fijne tijd aan de Trans na mijn zwangerschapsverlof. Wat moest ik wennen aan het nieuwe leven als moeder en aio tegelijk. Jullie hielpen mij loskomen van het moederschap om ook weer serieus aio te zijn, want dat proefschrift moest toch een keer afkomen. Gerdineke, jij kwam voor

Sanne in de plaats en werd stiekem een beetje 'mijn nieuwe Sanne'. Wat een kei ben jij! Als mens én als onderzoekster. Dankjewel voor al je hulp, voor de gezelligheid, voor je betrokkenheid en voor de altijd mooie verhalen en foto's. Ik ben heel blij en dankbaar dat jij mijn paranimf bent.

Anneloes, ondanks dat we nooit kamergenoten zijn geweest, voelde het wel zo. Dankjewel dat je altijd in was voor een praatje en met een brede lach voor me klaarstond op die maandagochtenden in de maanden na mijn verlof. Jij gaf mij het vertrouwen dat ik dat proefschrift écht wel af zou krijgen!

Aan allen die mijn leven nog mooier maken, bedankt!

Gelukkig is er ook nog een leven buiten de universiteit, dat net zo belangrijk (of misschien nog wel belangrijker) is voor het werk. Sharon, Ellen, Henriëtte en Jeroen, bedankt voor de gezellige avondjes waarin jullie mij op de hoogte hielden van het communicatievak in de praktijk. Shelly, Anneroos en Frerika, bedankt dat jullie al zolang mijn lieve vriendinnen zijn, bij wie ik altijd terecht kan – ik zou wel haast willen zeggen: in voor- en tegenspoed. Frerika, dank je wel dat je in die laatste maanden zoveel gezellige woensdagen hebt doorgebracht met Sterre, zodat ik nog even verder kon schrijven. Wetende dat Sterre het zo naar haar zin had met tante Freer, was het veel minder erg om op mijn 'vrije' dag te moeten werken.

Diane, dankjewel dat je me de kans hebt gegeven in de praktijk aan de slag te gaan met de expertise die ik in mijn onderzoek heb opgebouwd over het digitale zoekgedrag van kinderen. Dankjewel ook voor de vrijheid die je me hebt geboden om deze start in het bedrijfsleven te kunnen combineren met de laatste loodjes van mijn proefschrift.

Lieve ouders, schoonouders, zussen en zwagers, dank jullie wel dat jullie altijd geïnteresseerd en betrokken zijn geweest in mijn onderzoek en dat jullie vooral altijd héél trots op mij zijn. Dat maakt mij trots.

Lieve, lieve, Jaap, als ik jou in deze chronologie zou willen plaatsen, kwam je overal terug. Jij bent het begin en het einde. Dankjewel voor de ruimte die jij me gaf dit onderzoek te doen, voor je betrokkenheid, je geduld, je hulp bij het submitten van een artikel, al was het midden in de nacht, je opbeurende woorden als ik het even niet meer zag zitten, maar ook je schop onder mijn kont als dat nodig was. Dankjewel dat je mijn man bent en de papa van Sterre.

Lieve, lieve, lieve, Sterre! Ik draag dit boek op aan jou. Ik hoop dat ik met mijn onderzoek bij kan dragen aan een betere wereld voor jou; een wereld waarin digitale media zo zijn ontworpen dat jij als kind gemakkelijk informatie kunt zoeken en vinden.

SIKS Dissertation series

2014-04 Hanna Jochmann-Mannak (UT), Websites for children: Search strategies and interface design - Three studies on children's search performance and evaluation

2014-03 Sergio Raúl Duarte Torres (UT), Information Retrieval for Children: Search Behavior and Solutions

2014-02 Fiona Tulinayo (RUN), Combining System Dynamics with a Domain Modeling Method

2014-01 Nicola Barile (UU), Studies in Learning Monotone Models from Data

2013-43 Marc Bron (UVA), Exploration and Contextualization through Interaction and Concepts

2013-42 Léon Planken (TUD), Algorithms for Simple Temporal Reasoning

2013-41 Jochem Liem (UVA), Supporting the Conceptual Modelling of Dynamic Systems: A Knowledge Engineering Perspective on Qualitative Reasoning

2013-40 Pim Nijssen (UM), Monte-Carlo Tree Search for Multi-Player Games

2013-39 Joop de Jong (TUD), A Method for Enterprise Ontology based Design of Enterprise Information Systems

2013-38 Eelco den Heijer (VU), Autonomous Evolutionary Art

2013-37 Dirk Börner (OUN), Ambient Learning Displays

2013-36 Than Lam Hoang (TUe), Pattern Mining in Data Streams

2013-35 Abdallah El Ali (UvA), Minimal Mobile Human Computer Interaction

2013-34 Kien Tjin-Kam-Jet (UT), Distributed Deep Web Search

2013-33 Qi Gao (TUD), User Modeling and Personalization in the Microblogging Sphere

2013-32 Kamakshi Rajagopal (OUN), Networking For Learning; The role of Networking in a Lifelong Learner's Professional Development

2013-31 Dinh Khoa Nguyen (UvT), Blueprint Model and Language for Engineering Cloud Applications

2013-30 Joyce Nakatumba (TUE), Resource-Aware Business Process Management: Analysis and Support

2013-29 Iwan de Kok (UT), Listening Heads

2013-28 Frans van der Sluis (UT), When Complexity becomes Interesting: An Inquiry into the Information eXperience

2013-27 Mohammad Huq (UT), Inference-based Framework Managing Data Provenance

2013-26 Alireza Zarghami (UT), Architectural Support for Dynamic Homecare Service Provisioning

2013-25 Agnieszka Anna Latoszek-Berendsen (UM), Intention-based Decision Support. A new way of representing and implementing clinical guidelines in a Decision Support System

2013-24 Haitham Bou Ammar (UM), Automated Transfer in Reinforcement Learning

2013-23 Patricio de Alencar Silva (UvT), Value Activity Monitoring

2013-22 Tom Claassen (RUN), Causal Discovery and Logic

2013-21 Sander Wubben (UvT), Text-to-text generation by monolingual machine translation

2013-20 Katja Hofmann (UvA), Fast and Reliable Online Learning to Rank for Information Retrieval

2013-19 Renze Steenhuizen (TUD), Coordinated Multi-Agent Planning and Scheduling

2013-18 Jeroen Janssens (UvT), Outlier Selection and One-Class Classification

2013-17 Koen Kok (VU), The PowerMatcher: Smart Coordination for the Smart Electricity Grid

2013-16 Eric Kok (UU), Exploring the practical benefits of argumentation in multi-agent deliberation

2013-15 Daniel Hennes (UM), Multiagent Learning - Dynamic Games and Applications

2013-14 Jafar Tanha (UVA), Ensemble Approaches to Semi-Supervised Learning Learning

2013-13 Mohammad Safiri (UT), Service Tailoring: User-centric creation of integrated IT-based homecare services to support independent living of elderly

2013-12 Marian Razavian (VU), Knowledge-driven Migration to Services

2013-11 Evangelos Pournaras (TUD), Multi-level Reconfigurable Self-organization in Overlay Services

2013-10 Jeewanie Jayasinghe Arachchige (UvT), A Unified Modeling Framework for Service Design.

2013-09 Fabio Gori (RUN), Metagenomic Data Analysis: Computational Methods and Applications

2013-08 Robbert-Jan Merk (VU), Making enemies: cognitive modeling for opponent agents in fighter pilot simulators

2013-07 Giel van Lankveld (UvT), Quantifying Individual Player Differences

2013-06 Romulo Goncalves (CWI), The Data Cyclotron: Juggling Data and Queries for a Data Warehouse Audience

2013-05 Dulce Pumareja (UT), Groupware Requirements Evolutions Patterns

2013-04 Chetan Yadati (TUD), Coordinating autonomous planning and scheduling

2013-03 Szymon Klarman (VU), Reasoning with Contexts in Description Logics

2013-02 Erietta Liarou (CWI), MonetDB/DataCell: Leveraging the Column-store Database Technology for Efficient and Scalable Stream Processing

2013-01 Viorel Milea (EUR), News Analytics for Financial Decision Support

2012-51 Jeroen de Jong (TUD), Heuristics in Dynamic Sceduling; a practical framework with a case study in elevator dispatching

2012-50 Steven van Kervel (TUD), Ontologogy driven Enterprise Information Systems Engineering

2012-49 Michael Kaisers (UM), Learning against Learning - Evolutionary dynamics of reinforcement learning algorithms in strategic interactions

2012-48 Jorn Bakker (TUE), Handling Abrupt Changes in Evolving Time-series Data

2012-47 Manos Tsagkias (UVA), Mining Social Media: Tracking Content and Predicting Behavior

2012-46 Simon Carter (UVA), Exploration and Exploitation of Multilingual Data for Statistical Machine Translation

2012-45 Benedikt Kratz (UvT), A Model and Language for Business-aware Transactions

2012-44 Anna Tordai (VU), On Combining Alignment Techniques

2012-43 Withdrawn

2012-42 Dominique Verpoorten (OU), Reflection Amplifiers in self-regulated Learning

2012-41 Sebastian Kelle (OU), Game Design Patterns for Learning

2012-40 Agus Gunawan (UvT), Information Access for SMEs in Indonesia

2012-39 Hassan Fatemi (UT), Risk-aware design of value and coordination networks

2012-38 Selmar Smit (VU), Parameter Tuning and Scientific Testing in Evolutionary Algorithms

2012-37 Agnes Nakakawa (RUN), A Collaboration Process for Enterprise Architecture Creation

2012-36 Denis Ssebugwawo (RUN), Analysis and Evaluation of Collaborative Modeling Processes

2012-35 Evert Haasdijk (VU), Never Too Old To Learn -- On-line Evolution of Controllers in Swarm- and Modular Robotics

2012-34 Pavol Jancura (RUN), Evolutionary analysis in PPI networks and applications

2012-33 Rory Sie (OUN), Coalitions in Cooperation Networks (COCOON)

2012-32 Wietske Visser (TUD), Qualitative multi-criteria preference representation and reasoning

2012-31 Emily Bagarukayo (RUN), A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure

2012-30 Alina Pommeranz (TUD), Designing Human-Centered Systems for Reflective Decision Making

2012-29 Almer Tigelaar (UT), Peer-to-Peer Information Retrieval

2012-28 Nancy Pascall (UvT), Engendering Technology Empowering Women

2012-27 Hayrettin Gurkok (UT), Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games

2012-26 Emile de Maat (UVA), Making Sense of Legal Text

2012-25 Silja Eckartz (UT), Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application

2012-24 Laurens van der Werff (UT), Evaluation of Noisy Transcripts for Spoken Document Retrieval

2012-23 Christian Muehl (UT), Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction

2012-22 Thijs Vis (UvT), Intelligence, politie en veiligheidsdienst: verenigbare grootheden?

2012-21 Roberto Cornacchia (TUD), Querying Sparse Matrices for Information Retrieval

2012-20 Ali Bahramisharif (RUN), Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing

2012-19 Helen Schonenberg (TUE), What's Next? Operational Support for Business Process Execution

2012-18 Eltjo Poort (VU), Improving Solution Architecting Practices

2012-17 Amal Elgammal (UvT), Towards a Comprehensive Framework for Business Process Compliance

2012-16 Fiemke Both (VU), Helping people by understanding them - Ambient Agents supporting task execution and depression treatment

2012-15 Natalie van der Wal (VU), Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes.

2012-14 Evgeny Knutov (TUE) Generic Adaptation Framework for Unifying Adaptive Web-based Systems

2012-13 Suleman Shahid (UvT), Fun and Face: Exploring non-verbal expressions of emotion during playful interactions

2012-12 Kees van der Sluijs (TUE), Model Driven Design and Data Integration in Semantic Web Information Systems

2012-11 J.C.B. Rantham Prabhakara (TUE), Process Mining in the Large: Preprocessing, Discovery, and Diagnostics

2012-10 David Smits (TUE), Towards a Generic Distributed Adaptive Hypermedia Environment

2012-09 Ricardo Neisse (UT), Trust and Privacy Management Support for Context-Aware Service Platforms

2012-08 Gerben de Vries (UVA), Kernel Methods for Vessel Trajectories

2012-07 Rianne van Lambalgen (VU), When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions

2012-06 Wolfgang Reinhardt (OU), Awareness Support for Knowledge Workers in Research Networks

2012-05 Marijn Plomp (UU), Maturing Interorganisational Information Systems

2012-04 Jurriaan Souer (UU), Development of Content Management System-based Web Applications

2012-03 Adam Vanya (VU), Supporting Architecture Evolution by Mining Software Repositories

2012-02 Muhammad Umair (VU) Adaptivity, emotion, and Rationality in Human and Ambient Agent Models

2012-01 Terry Kakeeto (UvT), Relationship Marketing for SMEs in Uganda

2011-49 Andreea Niculescu (UT), Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality

2011-48 Mark Ter Maat (UT), Response Selection and Turn-taking for a Sensitive Artificial Listening Agent

2011-47 Azizi Bin Ab Aziz (VU), Exploring Computational Models for Intelligent Support of Persons with Depression

2011-46 Beibei Hu (TUD), Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work

2011-45 Herman Steehouwer (UvT), Statistical Language Models for Alternative Sequence Selection

2011-44 Boris Reuderink (UT), Robust Brain-Computer Interfaces

2011-43 Henk van der Schuur (UU), Process Improvement through Software Operation Knowledge

2011-42 Michal Sindlar (UU), Explaining Behavior through Mental State Attribution

2011-41 Luan Ibraimi (UT), Cryptographically Enforced Distributed Data Access Control

2011-40 Viktor Clerc (VU), Architectural Knowledge Management in Global Software Development

2011-39 Joost Westra (UU), Organizing Adaptation using Agents in Serious Games

2011-38 Nyree Lemmens (UM), Bee-inspired Distributed Optimization

2011-37 Adriana Burlutiu (RUN), Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference

2011-36 Erik van der Spek (UU), Experiments in serious game design: a cognitive approach

2011-35 Maaike Harbers (UU), Explaining Agent Behavior in Virtual Training

2011-34 Paolo Turrini (UU), Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations

2011-33 Tom van der Weide (UU), Arguing to Motivate Decisions

2011-32 Nees-Jan van Eck (EUR), Methodological Advances in Bibliometric Mapping of Science

2011-31 Ludo Waltman (EUR), Computational and Game-Theoretic Approaches for Modeling Bounded Rationality

2011-30 Egon van den Broek (UT), Affective Signal Processing (ASP): Unraveling the mystery of emotions

2011-29 Faisal Kamiran (TUE), Discrimination-aware Classification

2011-28 Rianne Kaptein (UVA), Effective Focused Retrieval by Exploiting Query Context and Document Structure

2011-27 Aniel Bhulai (VU), Dynamic website optimization through autonomous management of design patterns

2011-26 Matthijs Aart Pontier (VU), Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots

2011-25 Syed Waqar ul Qounain Jaffry (VU), Analysis and Validation of Models for Trust Dynamics

2011-24 Herwin van Welbergen (UT), Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior

2011-23 Wouter Weerkamp (UVA), Finding People and their Utterances in Social Media

2011-22 Junte Zhang (UVA), System Evaluation of Archival Description and Access

2011-21 Linda Terlouw (TUD), Modularization and Specification of Service-Oriented Systems

2011-20 Qing Gu (VU), Guiding service-oriented software engineering - A view-based approach

2011-19 Ellen Rusman (OU), The Mind ' s Eye on Personal Profiles

2011-18 Mark Ponsen (UM), Strategic Decision-Making in complex games

2011-17 Jiyin He (UVA), Exploring Topic Structure: Coherence, Diversity and Relatedness

2011-16 Maarten Schadd (UM), Selective Search in Games of Different Complexity

2011-15 Marijn Koolen (UvA), The Meaning of Structure: the Value of Link Evidence for Information Retrieval

2011-14 Milan Lovric (EUR), Behavioral Finance and Agent-Based Artificial Markets

2011-13 Xiaoyu Mao (UvT), Airport under Control. Multiagent Scheduling for Airport Ground Handling

2011-12 Carmen Bratosin (TUE), Grid Architecture for Distributed Process Mining

2011-11 Dhaval Vyas (UT), Designing for Awareness: An Experience-focused HCI Perspective

2011-10 Bart Bogaert (UvT), Cloud Content Contention

2011-09 Tim de Jong (OU), Contextualised Mobile Media for Learning

2011-08 Nieske Vergunst (UU), BDI-based Generation of Robust Task-Oriented Dialogues

2011-07 Yujia Cao (UT), Multimodal Information Presentation for High Load Human Computer Interaction

2011-06 Yiwen Wang (TUE), Semantically-Enhanced Recommendations in Cultural Heritage

2011-05 Base van der Raadt (VU), Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline

2011-04 Hado van Hasselt (UU), Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference learning algorithms

2011-03 Jan Martijn van der Werf (TUE), Compositional Design and Verification of Component-Based Information Systems

2011-02 Nick Tinnemeier (UU), Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language

2011-01 Botond Cseke (RUN), Variational Algorithms for Bayesian Inference in Latent Gaussian Models

2010-53 Edgar Meij (UVA), Combining Concepts and Language Models for Information Access

2010-52 Peter-Paul van Maanen (VU), Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention

2010-51 Alia Khairia Amin (CWI), Understanding and supporting information seeking tasks in multiple sources

2010-50 Bouke Huurnink (UVA), Search in Audiovisual Broadcast Archives

2010-49 Jahn-Takeshi Saito (UM), Solving difficult game positions

2010-48 Withdrawn

2010-47 Chen Li (UT), Mining Process Model Variants: Challenges, Techniques, Examples

2010-46 Vincent Pijpers (VU), e3alignment: Exploring Inter-Organizational Business-ICT Alignment

2010-45 Vasilios Andrikopoulos (UvT), A theory and model for the evolution of software services

2010-44 Pieter Bellekens (TUE), An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain

2010-43 Peter van Kranenburg (UU), A Computational Approach to Content-Based Retrieval of Folk Song Melodies

2010-42 Sybren de Kinderen (VU), Needs-driven service bundling in a multi-supplier setting - the computational e3-service approach

2010-41 Guillaume Chaslot (UM), Monte-Carlo Tree Search

2010-40 Mark van Assem (VU), Converting and Integrating Vocabularies for the Semantic Web

2010-39 Ghazanfar Farooq Siddiqui (VU), Integrative modeling of emotions in virtual agents

2010-38 Dirk Fahland (TUE), From Scenarios to components

2010-37 Niels Lohmann (TUE), Correctness of services and their composition

2010-36 Jose Janssen (OU), Paving the Way for Lifelong Learning; Facilitating competence development through a learning path specification

2010-35 Dolf Trieschnigg (UT), Proof of Concept: Concept-based Biomedical Information Retrieval

2010-34 Teduh Dirgahayu (UT), Interaction Design in Service Compositions

2010-33 Robin Aly (UT), Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval

2010-32 Marcel Hiel (UvT), An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems

2010-31 Victor de Boer (UVA), Ontology Enrichment from Heterogeneous Sources on the Web

2010-30 Marieke van Erp (UvT), Accessing Natural History - Discoveries in data cleaning, structuring, and retrieval

2010-29 Stratos Idreos (CWI), Database Cracking: Towards Auto-tuning Database Kernels

2010-28 Arne Koopman (UU), Characteristic Relational Patterns

2010-27 Marten Voulon (UL), Automatisch contracteren

2010-26 Ying Zhang (CWI), XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines

2010-25 Zulfiqar Ali Memon (VU), Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective

2010-24 Dmytro Tykhonov, Designing Generic and Efficient Negotiation Strategies

2010-23 Bas Steunebrink (UU), The Logical Structure of Emotions

2010-22 Michiel Hildebrand (CWI), End-user Support for Access to\\ Heterogeneous Linked Data

2010-21 Harold van Heerde (UT) Privacy-aware data management by means of data degradation

2010-20 Ivo Swartjes (UT), Whose Story Is It Anyway? How Improv Informs Agency and Authorship of Emergent Narrative

2010-19 Henriette Cramer (UvA), People's Responses to Autonomous and Adaptive Systems

2010-18 Charlotte Gerritsen (VU), Caught in the Act: Investigating Crime by Agent-Based Simulation

2010-17 Spyros Kotoulas (VU), Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications

2010-16 Sicco Verwer (TUD), Efficient Identification of Timed Automata, theory and practice

2010-15 Lianne Bodenstaff (UT), Managing Dependency Relations in Inter-Organizational Models

2010-14 Sander van Splunter (VU), Automated Web Service Reconfiguration

2010-13 Gianluigi Folino (RUN), High Performance Data Mining using Bio-inspired techniques

2010-12 Susan van den Braak (UU), Sensemaking software for crime analysis

2010-11 Adriaan Ter Mors (TUD), The world according to MARP: Multi-Agent Route Planning

2010-10 Rebecca Ong (UL), Mobile Communication and Protection of Children

2010-09 Hugo Kielman (UL), A Politiele gegevensverwerking en Privacy, Naar een effectieve waarborging

2010-08 Krzysztof Siewicz (UL), Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments

2010-07 Wim Fikkert (UT), Gesture interaction at a Distance

2010-06 Sander Bakkes (UvT), Rapid Adaptation of Video Game AI

2010-05 Claudia Hauff (UT), Predicting the Effectiveness of Queries and Retrieval Systems

2010-04 Olga Kulyk (UT), Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments

2010-03 Joost Geurts (CWI), A Document Engineering Model and Processing Framework for Multimedia documents

2010-02 Ingo Wassink (UT), Work flows in Life Science

2010-01 Matthijs van Leeuwen (UU), Patterns that Matter

2009-46 Loredana Afanasiev (UvA), Querying XML: Benchmarks and Recursion

2009-45 Jilles Vreeken (UU), Making Pattern Mining Useful

2009-44 Roberto Santana Tapia (UT), Assessing Business-IT Alignment in Networked Organizations

2009-43 Virginia Nunes Leal Franqueira (UT), Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients

2009-42 Toine Bogers (UvT), Recommender Systems for Social Bookmarking

2009-41 Igor Berezhnyy (UvT), Digital Analysis of Paintings

2009-40 Stephan Raaijmakers (UvT), Multinomial Language Learning: Investigations into the Geometry of Language

2009-39 Christian Stahl (TUE, Humboldt-Universitaet zu Berlin), Service Substitution -- A Behavioral Approach Based on Petri Nets

2009-38 Riina Vuorikari (OU), Tags and self-organisation: a metadata ecology for learning resources in a multilingual context

2009-37 Hendrik Drachsler (OUN), Navigation Support for Learners in Informal Learning Networks

2009-36 Marco Kalz (OUN), Placement Support for Learners in Learning Networks

2009-35 Wouter Koelewijn (UL), Privacy en Politiegegevens; Over geautomatiseerde normatieve informatie-uitwisseling

2009-34 Inge van de Weerd (UU), Advancing in Software Product Management: An Incremental Method Engineering Approach

2009-33 Khiet Truong (UT), How Does Real Affect Affect Affect Recognition In Speech?

2009-32 Rik Farenhorst (VU) and Remco de Boer (VU), Architectural Knowledge Management: Supporting Architects and Auditors

2009-31 Sofiya Katrenko (UVA), A Closer Look at Learning Relations from Text

2009-30 Marcin Zukowski (CWI), Balancing vectorized query execution with bandwidth-optimized storage

2009-29 Stanislav Pokraev (UT), Model-Driven Semantic Integration of Service-Oriented Applications

2009-28 Sander Evers (UT), Sensor Data Management with Probabilistic Models

2009-27 Christian Glahn (OU), Contextual Support of social Engagement and Reflection on the Web

2009-26 Fernando Koch (UU), An Agent-Based Model for the Development of Intelligent Mobile Services

2009-25 Alex van Ballegooij (CWI), "RAM: Array Database Management through Relational Mapping"

2009-24 Annerieke Heuvelink (VUA), Cognitive Models for Training Simulations

2009-23 Peter Hofgesang (VU), Modelling Web Usage in a Changing Environment

2009-22 Pavel Serdyukov (UT), Search For Expertise: Going beyond direct evidence

2009-21 Stijn Vanderlooy (UM), Ranking and Reliable Classification

2009-20 Bob van der Vecht (UU), Adjustable Autonomy: Controling Influences on Decision Making

2009-19 Valentin Robu (CWI), Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets

2009-18 Fabian Groffen (CWI), Armada, An Evolving Database System

2009-17 Laurens van der Maaten (UvT), Feature Extraction from Visual Data

2009-16 Fritz Reul (UvT), New Architectures in Computer Chess

2009-15 Rinke Hoekstra (UVA), Ontology Representation - Design Patterns and Ontologies that Make Sense

2009-14 Maksym Korotkiy (VU), From ontology-enabled services to service-enabled ontologies (making ontologies work in e-science with ONTO-SOA)

2009-13 Steven de Jong (UM), Fairness in Multi-Agent Systems

2009-12 Peter Massuthe (TUE, Humboldt-Universitaet zu Berlin), Operating Guidelines for Services

2009-11 Alexander Boer (UVA), Legal Theory, Sources of Law & the Semantic Web

2009-10 Jan Wielemaker (UVA), Logic programming for knowledge-intensive interactive applications

2009-09 Benjamin Kanagwa (RUN), Design, Discovery and Construction of Service-oriented Systems

2009-08 Volker Nannen (VU), Evolutionary Agent-Based Policy Analysis in Dynamic Environments

2009-07 Ronald Poppe (UT), Discriminative Vision-Based Recovery and Recognition of Human Motion

2009-06 Muhammad Subianto (UU), Understanding Classification

2009-05 Sietse Overbeek (RUN), Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality

2009-04 Josephine Nabukenya (RUN), Improving the Quality of Organisational Policy Making using Collaboration Engineering

2009-03 Hans Stol (UvT), A Framework for Evidence-based Policy Making Using IT

2009-02 Willem Robert van Hage (VU), Evaluating Ontology-Alignment Techniques

2009-01 Rasa Jurgelenaite (RUN), Symmetric Causal Independence Models