Analysis of Search and Browsing Behavior of Young Users on the Web

Sergio Duarte Torres [*], University of Twente
Ingmar Weber [*], Qatar Computing Research Institute
Djoerd Hiemstra, University of Twente

The Internet is increasingly used by young children for all kinds of purposes. Nonetheless, there are not many resources especially designed for children on the Internet and most of the content online is designed for grown up users. This situation is problematic if we consider the large differences between young users and adults since their topic interests, computer skills and language capabilities evolve rapidly during the childhood. There is little research aimed at exploring and measuring the difficulties that children encounter on the Internet when searching for information and browsing for content. In the first part of this work, we employed query logs from a commercial search engine to quantify the difficulties children of different ages encounter on the Internet and to characterize the topics that they search for. We employed query metrics (e.g. the fraction of queries posed in natural language), session metrics (e.g. the fraction of abandoned sessions) and click activity (e.g. the fraction of ad clicks). The search logs were also used to retrace stages of children development. Concretely, we looked for changes in interests (e.g. the distribution of topics searched) and language development (e.g. the readability of the content accessed and the vocabulary size).

In the second part of this work, we employed toolbar logs from a commercial search engine to characterize the browsing behavior of young users, particularly to understand the activities on the Internet that trigger search. We quantified the proportion of browsing and search activity in the toolbar sessions and we estimated the likelihood of a user to carry out search on the Web vertical and multimedia verticals (i.e. videos and images) given that the previous event is another search event or a browsing event.

We observed that these metrics clearly demonstrate an increased level of confusion and unsuccessful search sessions among children. We also found a clear relation between the reading level of the clicked pages and characteristics of the users such as age and educational attainment.

In terms of browsing behavior, children were found to start their activities on the Internet with a search engine (instead of directly browsing content) more often than adults. We also observed a significantly larger amount of browsing activity for the case of teenager users. Interestingly we also found that if children visit knowledge related websites (i.e. information-dense pages such as Wikipedia articles), they subsequently do more web searches than adults. Additionally, children and especially teenagers were found to have a greater tendency to engage in multimedia search, which calls to improve the aggregation of multimedia results into the current search result pages.

Categories and Subject Descriptors: H.3.3 [**Information Search and Retrieval**]: Query formulation, Search process

General Terms: Experimentation, Human-factors, Performance

Additional Key Words and Phrases: children, query logs, toolbar logs, search behavior, browsing behavior, session analysis, topic classification, web search, young adults, adults, Yahoo! Search, Yahoo! Answers.

---

1. INTRODUCTION

The fraction of children using the Web and the amount of time they spend online has
increased significantly in the past years. A case study carried out in 2008 involving up
to 2,500 in-home interviews with children and their parents in the UK, reported that
63% and 76% of users aged 5 to 7 and 8 to 11 years old, respectively, use the Internet
at home [Child Trends Data Bank 2013]. In the US, 32.4 million of children under the
age of 18 years old were active users of the Internet in the same year, accounting for up
to 19% of the online population. Similar trends have been reported in other developed
countries [Ofcom 2010].

More recent studies carried out in the European Union have pointed out that not
only the access to the Internet continues to increase among the young population but
also the amount of time they spend online. Livingstone et al. [2011] reported from a
detailed survey carried out in 2009 to 2011 with European children and their parents
in 25 countries, that users from 9 to 16 year old spend on average 88 minutes per day
online. They also found that 33% of these users go online via mobile phones, 87% at
home, and 49% at home and from their bedroom. Even higher Internet access percent-
ages have been reported in the last years for this segment of users in the US.

Madden et al. [2013] found through a survey conducted in 2012 with 802 parents
and their teenagers aged 12 to 17 years old, that 95% of the teenagers use the Inter-
net regularly, 78% own a cell phone and around 47% own a smart-phone, which is a
prominent mean to access the Internet. It has also been shown that children are often
trusted to search the Internet on their own, 68% and 84% for children aged 5-7 and
8-15 in the UK, respectively [Ofcom 2010]. Undoubtedly, the access and use of the In-
ternet by children will keep increasing in these and other regions of the world in the
coming years.

Given the small amount of content carefully designed for this audience and the lack
of specialized search engines dedicated to help children to find appropriate content on
the Web, there is an increasing need for research aimed at understanding the activities
that these users engaged on the Internet.

The motivation of this study is to characterize the search behavior of young users
and the browsing activities that lead to search. In terms of search behavior, we focus
on identifying the difficulties that young users encounter on the Internet when they
search for information with a state of the art search engine. Given that current search
engines are not designed for young users, we hypothesize that they struggle in their
way to find information on the web and that these difficulties are quantifiable from
their interactions with the search engine. The interactions are registered in the logs of
the search engine. In this study we employed a large scale log sample from the Yahoo!
search engine and the Yahoo! toolbar application. Identifying the problems young users
face when searching information on the web, along their topics of interest, is key in the
development of search and web services tailored at children and teenagers.

The difficulties that children encounter during the search process are exemplified
through the following two search sessions derived from the log sample studied: *(1)* A
10 years old girl submits the query *what is love*, the search engine triggers advertise-
ments related to dating and casual encounters. Thinking that this ad is a result to the
query, the girl clicks on it, after spending few seconds trying to understand what is
happening, the user goes back and then clicks on the first web result, which explains
the chemical processes involved when people feel love. The content of this website goes

beyond her reading skills and the user quits the search session, most likely dissatisfied; *(2)* When a 9 years old boy submits the query *hun*, the search engine suggests queries such as hun school (Princeton college), hun sen (primer minister of Cambodia) and hun empire (former empire ruled by Attila). Although this user is probably targeting the last topic suggested by the search engine, the user does not seem to notice any of the query suggestions and simply continues with the initial query. Then, the user clicks on the first web result, which happens to be a web directory of links with adult content. Hun is also a popular term used to refer to a specific type of adult content. The user, who is probably confused by the content, decides to go back, then the user clicks on the second web result, which is the Wikipedia article of the Hun empire. As it was the case with our previous example, the article is dense and its language is too advanced for the reading capabilities of this user, who after few seconds aborts the search session. In these two examples, we observed that young users have a tendency to click on higher ranked results, spend short times on each url, and in general have shorter sessions than those observed in older users.

The first contribution of this paper is to identify and quantify the search difficulties of young users using search engines and characterize the topics they are interested in, by using a large query log sample. We are also interested in contrasting the search difficulties and topic interests of young and adult users. At the best of our knowledge, this paper represents the first research effort to investigate on large scale the search behavior of children and teenagers and to identify unobtrusively the problems that they encounter with an state-of-the-art search engine. Even though several valuable case studies with children users have been reported [Bilal 2000; 2001; 2002; Druin et al. 2009a], these studies are highly obtrusive and include a limited number of users and a small number, which are assigned a limited number of artificial information tasks. These research studies are also unable to capture a representative overview of the broad spectrum of topics that motivate web search in the young population. In the same lime, we are interested in exploring signs of development stages through the usage of a search engine. For instance, to trace the evolution in terms of the readability of the content clicked and the sentiment reflected in the queries.

The second contribution of this paper is to characterize the type of activities that young users carried out on the Internet when a search engine is not being used, and to understand which of these activities are more likely to motivate search. The understanding of the online activities engaged by young users and how they relate to their search queries leads to an integral picture of search engine usage. At the best of our knowledge no research has been carried out in this regard that is both: on a large scale and unobtrusive. Similarly, we believe this is the first research attempt to relate on a large scale, the search and browsing behavior of young users. It is important to mention that recent research has addressed the characterizing of browsing activities engaged by the average web user [Kumar and Tomkins 2010; Cheng et al. 2010], however these studies are not oriented towards the young population and in general the focus is not on the age demographic dimension of the users.

1.1. Research questions

The contribution of this paper can be summarized by the following research questions:

— **RQ-1**: Do young users struggle to find information with a large scale search engine, and how is this struggle reflected in their search behavior from a query log perspective?
— **RQ-2**: Do the search behavior and search difficulties of children, teenagers and adults differ in a large scale search engine (Yahoo! Search)?

—**RQ-3**: Can we retrace stages of children and teenager development in terms of the topics they are interested in through their queries and the characteristics of these queries?

—**RQ-4**: What other activities are carried out by young users on the Web browser? How prominent is browsing for each age range? At what ages is multimedia search preferred?

—**RQ-5**: Which type of search and browsing activities are more likely to trigger search in web and multimedia search engines in the case of young users? To what extent do these triggers differ from those observed in adults users?

We tackle *RQ-1* to *RQ-3* by using a large query log sample from the Yahoo! search engine. The search logs registered queries and search activity of users from 7 to 70 years old. The logs were taken from the US market in a time window of 4 months in 2010. The methods employed in this study consist of using well established query log metrics and novel metrics targeted at understanding the search behavior and quantifying the search problems that children encounter when searching information on the Web.

In respect to *RQ-1* we hypothesize that young users (and particularly users up to 12 years old) have greater difficulties than adults to search for information on the web, and that this behavior is manifested in the way young users construct their queries and explore the result list. Particularly, we believe young users have problems specifying their information needs with keywords, and that young users make more use of natural language in their queries. In terms of result exploration, we expect young users to quit their search sessions sooner, and to have shorter clicks on results than adult as a consequence of their search frustration. We also expect to observe greater use of the same queries and visited results previously clicked within the search session, which has been observed in the literature [Bilal 2000; 2002]. Lower usage of query suggestion is expected since current search engines suggest terms for all user ages, thus most query suggestions are likely to be unsuitable for children.

We collect evidence to address these hypotheses through the following query log metrics: At a query level, we include average query length (measure of query specificity), query structure and natural language usage in queries. At a click level, we considered the click position and click duration distribution to analyze the way young users explore results. At the session level we explored session duration, session length, query re-finding, click re-finding, query suggestion and query correction usage. Further clarifications on how these metrics provide evidence of search difficulty will be given when reporting each result.

It is important to mention that after manually inspecting a small sample of search sessions, clicks on adult content and on advertisement at very young ages were observed. Motivated by this observation, we formally estimated the likelihood of accidentally click on adult content, and the likelihood of clicking on advertisement for young and adult users. These query log metrics are novel and oriented to collect evidence of the problems children encounter on the Web, since both type of content (adult content and ads) are clearly not targeted at children. High volumes of these clicks suggest search disorientation and difficulties to access appropriate information. We hypothesize that young children click on adult content by mistake more often than adults, and that they have a comparable proportion of clicks on ads in respect to other age groups given their difficulty to identify this type of content as advertisement.

RQ-3 is addressed by pointing out differences in the topic distribution of the queries submitted by users of different ages. The topic distribution is extracted by classifying the queries submitted to the Yahoo! search engine using the Yahoo! Directory categories. We also employed a fine-grained topic classification using the categories of the

Yahoo! Answers system, which allows to have a more detailed overview of the topics associated to the questions and concerns expressed in the children and teenagers queries. It is important to clarify that we are not analyzing the queries submitted to the Yahoo! Directory and Yahoo! Answers service, we solely used the category hierarchy provided by these two services to map the queries submitted to the Yahoo! search engine to topics.

We also show that the development stages are reflected in the search logs through the reading level of the pages clicked. As an aside, we employ other demographics factors such as the average income of the users location to show its influence in the reading level of the Web results clicked.

The vocabulary size is another important feature of children development. Specifically, we measure the vocabulary size by counting the average number of unique words in the queries. We carry out an analogous procedure to find the *web resource* vocabulary of the users ,which is defined as the number of distinct *urls* and domains accessed by users in a given age range. As a last metric, the sentiment expressed in the queries is quantified. We expect greater usage of sentiment at younger ages.

We expect to find clear differences in terms of the topic distribution of children, teenagers and adults, and in general we expect to see a correlation between age and topic distribution. Similarly, we expect a higher proportion of clicks on basic language content and a smaller vocabulary size at younger ages given that children have less developed vocabulary and language capabilities. All the query metrics mentioned so far are also employed to address *RQ-2*. This is carried out by contrasting the results across age groups. For instance, we can characterize the way users submit queries by comparing the characteristics (e.g. query length, natural language usage) between users from 7 to 12 (children), 13 to 18 (teenagers) and above 18 (adults).

Analogously, research questions *RQ-4* and *RQ-5* are addressed by analysing a set of carefully chosen log metrics estimated on a per-user basis. However, these two research questions are addressed using a large sample from the Yahoo! Toolbar logs instead of the Yahoo! Search logs. The toolbar logs contain records of the urls accessed in a browser, which may include search activity and browsing activity (e.g. emailing, social networking websites). All the toolbar data is captured only after the user has explicitly opted-in to have their browsing activity registered. All the data employed in our studies had been previously anonymized. This data allows to analyze the online activity of users of different ages outside the search engine, including the activity before they search the web. The log data employed corresponds to the same market and time window of the query logs employed to address *RQ-1* to *RQ-3*.

In this work, we refer to browsing activity as all the action that is not carried out within the scope of the search engine. Thus, with *browsing* we refer to website visits such as email portals, social networking sites (e.g. Facebook), news portals (e.g. BBC), head-listing websites (e.g. Ebay) and in general any website that is accessed directly by typing the url in the browser or through a bookmark. Note that this differs from the *search by browsing* definition addressed in previous case-studies of children's search behavior [Druin et al. 2009a], in which children are asked to engage informational tasks under two settings: keyword search (i.e. search engines) and by browsing a hierarchy of categories. For us *browsing activity* refers to browsing the Internet in general and not necessarily to search for information through a hierarchy of categories.

RQ-4 is addressed by estimating the volume of search and browsing activity for all the age ranges. We analyze the ratio between search and browsing activity to contrast how users of different ages spend their time on the Internet. A more detailed understanding of the browsing activities is obtained by estimating the likelihood of a user to submit a query in a search engine given that their previous activity was a browsing activity. Particularly we analyze web search and *multimedia* search separately.

Another advantage of the toolbar logs is that access to multimedia verticals such as Yahoo! Images and Yahoo! Videos are also registered. We hypothesize that young users browse more than they search because non-young users have a greater expertise with search engines, which is likely to be reflected in the amount of search that is carried out by these users. We also hypothesize that multimedia search is preferred by young audiences given that the rich content found in these Web services is more engaging.

RQ-5 is addressed by analysing the likelihood of web and multimedia search to trigger other searches. A more detailed analysis of the likelihood of carrying out search after a browsing event is carried out to address this research question. Concretely, we brake down the cases in which the search query is explicitly mentioned in the browsing website, the cases in which it is mentioned in its domain, and the cases in which it is not mentioned at all. This analysis provides direct evidence of the specific types of browsing activity that lead to query submissions in a search engine. For instance, if the query is mentioned in the browsing website that was clicked prior to the query submission,

We hypothesize that search in each age range is motivated by specific type of browsing events. For instance, we expect knowledge intense websites (e.g. Wikipedia) to be more likely to trigger searches of young users (e.g. children) than of adult users. Similarly, we expect social networking page views to trigger search more frequently in teenagers.

1.2. Paper Organization

The paper is organized as follows: in Section 2 we present the most relevant related work of previous studies on query logs and on children's search behavior. We also provide the most related literature of toolbar logs mining. Section 3 describes the research method that is followed in this research paper. In Section 3, the data set used is described and the cleaning steps are explained. In Section 4, we present the results and discussion of children's search difficult (although few of items shown in this section also support *RQ-3*. Concretely we report results regarding query structure, click distribution, click duration, session characteristics and query assistance usage. Section 5 discusses our findings on retracing child development stages using query logs. These include topic distribution description, vocabulary size, query sentiment analysis and readability of the content clicked. Section 6 describes the Yahoo! toolbar sample employed in this study and data cleaning steps. A description of the method to extract and analyze measures from this data is also addressed. Particularly, we define the type of events that are considered for this part of the study.

Section 7 presents the discussion regarding the session characteristics found in the toolbar logs. These characteristics are contrasts along the age demographic dimension (this refers to *RQ-4*). Section 8 discusses the likelihoods of different browsing pageview and search activity that lead to web and multimedia search. In Section 9, we expand the analysis carried out in Section 8 by quantifying the cases in which the query is mentioned in the url and domain previously browsed. Sections 8 and 9 addressed *RQ-4* and *RQ-5*. Finally, section 10 concludes this work with a discussion of our main findings and how they could be applied to state of the art search engines. We also provide recommendations for future work.

1.3. Limitations of this study

In this study, we focus on understanding the search and browse behavior of young users (children and teenagers). The findings are contrasted against the search behavior of adult users. However, the analysis of the difference within adult users of different ages (e.g. young adults and seniors) is out of the scope of this study. Little research has been dedicated to understand the search behavior and difficulties that senior users en-

counter on the Internet. We believe an exclusive line of research is required to address this.

The results presented in this study are derived from log data from the U.S. market in which the predominant language is English. Further research is needed to address cultural differences in search behavior and in pinpointing the search difficulties of users from different cultural origins and different languages. The evidence collected throughout this study focus on search and browsing metrics captured by means of log data. Other aspects of the content explored by the users, such as the layout characteristics of the websites, font types, sizes and other features that are not reflected in the query logs are not addressed and are out of the scope of this study. Further research is needed to relate aesthetic and functional features of websites with the search behavior of users.

2. RELATED WORK

The most relevant literature on children search behavior, query log analysis, browsing behavior and information retrieval for children are described in the following paragraphs.

2.1. Information seeking by children

The first studies attempting to characterize the search behavior of children have been carried out using non-internet systems, such as electronic libraries, CD-ROMs and OPACs (Online Public Access Catalogs). Solomon [1993] explored the search success of elementary school children when using an OPACs. The authors found that children were able to use the system effectively when engaging simple searches. However, they found that complex searches were hampered by the lack of mechanical skills of children. They pointed out that factors such as typing on the keyboard, spelling, limited vocabulary and reading expertise are skills that are not developed enough in children in order to use the OPAC system studied [Broch 2000]. Borgman et al. [1995] found a similar behavior with high school children and a different OPAC. They also reported that these children had conceptual difficulties categorizing and browsing for searches that are more complex. Similarly, Neuman [1995] found from a survey including 25 digital library administrators, that the main problems children encountered during the search on digital libraries are the generation of keywords to construct the query, and the lack of effective search strategies.

Recent studies have explored the search behavior of children on the Internet with search engines. Nahl and Harada [1996] carried out a study with 191 high school students to determine their search effectiveness after they have received special training to search the Internet. Users were asked to solve specific information tasks on the Internet. They were assessed based on the information they collected. Nahl and Harada [1996] reported that most of the students had difficulties understanding how the search query is constructed with boolean and default operators. In this study was also observed that the lack of adequate vocabulary and content knowledge led to difficulties in the search process.

Bilal and Watson [1998] conducted a case study with children from a 7th grade science class (children between 11 and 13 years old) to determine how this group of users solve frequent school information tasks on the Web directory Yahooligans![1]. This web service provides a directory structure in which users can browse from a large collection of websites. A search box is also provided to let users formulate search queries to find websites matching the query terms. Bilal and Watson [1998] found that children tend to ignore the browsing categories and that they start their search directly

---

[1]Today known as Yahoo! Kids: http://kids.yahoo.com/

using the search box utility. In the search box the mechanical problems identified in the previous research on digital libraries were also observed [Nahl and Harada 1996; Broch 2000]. The search effectiveness was hampered by the misspelling problem of the users, the lack of understanding in the use of logical operators and the formulation of queries using natural language, which were not treated adequately by the search services studied. It was also pointed out that certain queries lead only to a small amount of appropriate content for the age of the users.

Bilal [2000; Bilal [2002] carried out a follow up study of search behavior and usage of Yahooligans! with a sample of 17 users from 11 to 13 years old. Children were asked to solve open and well-defined informational tasks under two scenarios: informational task designed by the researchers and self informational tasks in which children were allowed to freely conduct their searches. In general, children were found to have limited success with the tasks given their lack of developed search skills and mechanical problems. Children also had trouble selecting the right categories in Yahooligans! In terms of browsing behavior children rarely explored thoroughly the results returned, they were found to have a search looping behavior in which previous seen results were often accessed again, and the back button of the browser was frequently activated. The authors also observed a lack of engagement when carrying out well established tasks, which also hampered their search experience. On overall, children lacked of focus and seemed disoriented during the search process. The authors also pointed out that the design of Yahooligans! is not well suited for children of the age studied. Children were found to perform better under the second scenario (self-assigned tasks), in which they showed a higher tendency to engage a navigational approach, instead of a keyword search approach. The authors argued that this occurred due to the poor keyword search capabilities of the system and the greater engagement level of children when they define their own search goals.

Bowler et al. [2001] studied the search process of a small group of children aged 11 to 12 to solve school informational tasks on the Internet. The authors reported that the search engines employed (Excite, AltaVista and Yahoo!) contained information for all audiences, which discourage the users since they took long time periods to find useful pieces of information. The authors argued that the overwhelming volume of information delivered for each query led users to dead ends and the visit of previously accessed link. Additionally, children were observed to trust blindly in the results returned by the search engine making more difficult for children to assess the quality of the results.

Druin et al. [2010; Druin et al. [2009a] characterized the search roles that children (aged 7 to 11 years old) adopt during the search process and studied how these roles depend on the children's environment and their motivation. They found that the computer expertise and orientation to explore visual content varies not only between children but also within the type of the information task. Kammerer and Bohnacker [2012] found similar trends in a recent study with 21 children aged 8 to 10. In their study, children were asked to engage informational tasks in the Google search engine. They found that children only used few keywords, which often led to an ambiguous set of web results mixing content that is suitable and non-suitable for children. They also observed that the search performance improved when using queries that are more specific.

Fidel et al. [1999] conducted a case study with eight teenager users (aged 16 to 18) in a high school library. Users were giving school assignments to be solved with a search engine. No special training was provided for the task. No restrictions were established for the search engine, and users were allowed to use their favorite search tool. Most the students opted blindly for the search engine automatically adopted by the Internet browser. The users in this age range also ignored the category browsing functionality of some search engines and favored the submission of queries. Nonetheless, these users

were found to perform poorly when searching for information and they were found to reuse keywords, to have poor spelling in the formulation of queries and to revisit previous website,s even if they were off-topic for the search task. In the same line, Gunn and Hepburn [2003] observed the search information strategies and general usage of search engines of twelfth grade students (users aged 17 to 18 years old). They found a mismatch between the self-perception of search skills by these users and their actual performance in finding quality information. The users reported themselves as good web searchers, however they were unaware of the usage of boolean operators and other mechanism to refine the search. Surprisingly most users were also unaware of search engines mechanism to narrow the search to other media type such as images.

Jochmann-Mannak et al. [2010; Jochmann-Mannak et al. [2012] evaluated the preferences of children towards web pages with several layouts designed for children. The authors also compared search engines designed for children against Google. The case study was carried out with a group of 32 children between 8 to 12 years old. Surprisingly, they found that children tended to prefer the Google like interface to carry out their searches on the Internet. The authors found that browsing interfaces designed around children metaphors were not of the like of these users and that in rare cases these interfaces added value to a Google like service.

On overall, most of the previous studies discuss the search problems caused by the mechanical and cognitive skills of children when searching on the Internet and the mismatched between current search engines and children search capabilities. Even systems and websites that are aimed at children were not satisfying when assigning specific search tasks, as it was the case with Yahooligans! [Bilal 2000; 2002]. Opposite results have been shown in terms of the search approached preferred by children between 8 to 12 years old. In Yahooligans! it was found that browsing search was preferred over keyword search [Bilal 2002] , however in most of the other studies the opposite was reported [Jochmann-Mannak et al. 2010; Jochmann-Mannak et al. 2012]. From the results of these studies, it seems that even though children tend to prefer the keyword search environment, they performed better on the browsing style search given that these systems mitigate the mechanical skills of children towards spelling and query formulation.

It is important to mention that all the studies mentioned so far consider only a small group of users and focus on a specific age range. Our work differs from theirs in that we quantify the search characteristics and search difficulty of children based on aggregated results of thousands of users across a broad age range, unobtrusively, which makes our observations more representative on a web-scale. Additionally we report topic interest trends over a population with diverse demographic characteristics, which is not possible to observed with a limited number of users. Moreover, no study mentioned so far contrasted the search behavior between age ranges, which we address in this chapter using fine-grained age ranges of young users. Another important research concern that is not addressed in any of the studies mentioned is the understanding of the activities that children carry out on the Internet browser outside the scope of a search engine, and how these activities motivate search in young users. We address this research gap through the analysis of the Yahoo! toolbar logs.

2.2. Related query log analysis

Duarte Torres et al. [2010; Torres et al. [2010] constructed two set of search sessions from the AOL search logs: the first with users accessing information aimed at the general public (i.e. content targeted at non-children users), and the second with users accessing information aimed at children. The latter set of search sessions were constructed by using a set of carefully selected urls aimed at children from the Kids and Teens section of the Dmoz Open Directory. The search behavior of the users in both

search sessions were compared. Although it is not possible to ensure that the users studied in the latter set were children, significant differences between the two sets were found. Some of the results are in line with previous case studies on children search behavior. For instance, a greater usage of natural language and evidence of loopy behavior was found in the set of queries aimed at content for children. This behavior has been reported before [Bilal 2001; Fidel et al. 1999; Bilal and Watson 1998].

Duarte Torres and Weber [2011] characterized and contrast the behavior of children, teenager and adults by using a large search log sample of users of different age ranges. The age ranges represent stages of human development and they were obtained by using the birth year reported in the user profile information. Several query log metrics at the query and session level were employed to quantify the search difficulty that young users experience on the Internet. A broad set of topics searched by children, teenagers and adults was identified. We found a prominent click bias towards top ranked pages by young users, a smaller query length average on younger users, greater usage of natural language by teenagers and we provided preliminary results in terms of the topic interests of young users and their contrast against adult. In this paper we greatly expand on these results by adding more depth to the topic interest of users and understanding the browser activities of young and adult users.

Weber and Castillo [2010] presented a query logs study on how search differs on users with different demographics. They used demographic information derived from the US-census and user profile information to describe search patterns and behaviors for population segments with different demographic characteristics. In this chapter, we employed an analogous methodology to show that the reading level of the urls clicked by children also varies across demographic features.

Weber and Jaimes [2011] studied the relations between the dimensions "who searches", "what they search", and the "how they search" interact. Related to our work, they also gave details about topical distributions as function of the user's age. Even though they provided detailed topic results for adult users, the young user topic characteristics are very broad since they aggregated the logs of users from 7 to 25 years old into a single group.

In our study we apply similar methodological techniques, such as analyzing session characteristic, however the main difference is one of breadth against depth, in which our focus is on a particular demographic group, namely young users (children and teenagers). In this way, we provide more detailed results regarding the search behavior and difficulties that young users face when searching for information online.

Beheshti et al. [2010] performed an analysis of transaction logs from the portal *History trek*, which contains Canadian historic educational content for children. The logs analyzed contain up to 92K transactions and it was registered from 2007 to 2009. The authors found that the hierarchy of categories and subjects provided in the interface of the portal accounted for 83% of the searches, which indicates a clear preference for browsing over keyword based search. Beheshti et al. [2010] pointed out that this behavior is potentially due to the greater cognitive load associated to the formulation of search queries in respect to the browsing of predefined categories. Even though this result seems contradictory to some of the findings of Bilal [2000; Bilal [2002] and Fidel et al. [1999], it is important to recall that Beheshti et al. [2010] employed a carefully designed interface aimed at children. contrary to the first studies which were carried out on general purpose search engines.

More recent query log analysis have been carried out to understand the behavior of children when using search engines that have been adapted for them, in terms of search interface and content. Gossen et al. [2011] showed preliminary results from a query log analysis of three German search engines providing only content suitable for children. They found that children tend to submit shorter queries than adults (which is

also found in our study) and that, as it is expected, children commit a greater number of spelling mistakes.

2.3. Browsing behavior

Kumar and Tomkins [2010] carried out a large-scale user behavior study on search and toolbar data using logs from the Yahoo! system. They studied search sessions of adult users, in general the age dimension was not addressed in their study. They proposed a classification of browsing page views based on content (e.g. news games, portals), communication (e.g. email, social networking) and search (e.g. web, multimedia). They found that around half of the page views belong to the content categories and a third to the communication category. The chain of referrals (i.e. the chain of pages visited on the Internet browser) was analyzed to characterize the way users navigate through the page views. They showed that mail, news and social bookmarking page views occur in isolation and that on overall 35% did not have any referral. In this paper, we employ a similar classification of browsing page views. Our work differs from the work of Kumar and Tomkins [2010], in that we analyze the browsing behavior of users not only based on page view content classification but also according the age of the user.

Cheng et al. [2010] presented a method to predict the search intent of users by considering their browsing behavior. They first carried out a user study to characterize the main types of page views that lead to web searches by using toolbar data. Then, they employed a machine learning approach to rank queries according the previous browsing activities of the user. In this paper we are interested in understanding which type of browsing activities lead to web and multimedia search and how these activities differ with the age of the user.

Goel et al. [2012] carried out a large scale study of browsing behavior of users of different demographic characteristics by employing web panel data and user-level demographics such as age, sex, race, education and income. In terms of age their study included users from 7 years old to 80 years old. They found that all demographic groups spend the majority of their time online on the same type of activities: social media and emailing. In this regard, they observed that teenagers aged 15 use the most social media and its usage decreases consistently after this age. On the other hand, they observed that emailing usage correlates positively with age. They also found pronounced differences regarding the frequency in which different demographic groups (in terms of salary income and education) access websites with news, health and reference (encyclopedia) content. They also explored the digital divide in terms of Internet usage for research and informational queries, its relation to the educational background of the users and other demographic features such as salary income, race and gender. Our work differs from theirs in that we focus on identifying the browsing activities of very young users. We also emphasize our analysis in the browsing activities that motive search.

On overall, the previous research on browsing behavior on the Internet has focused on providing a framework to classify the possible page views, and to provide automatic methods to predict a page view type given the previous pages accessed by the user [Kumar and Tomkins 2010; Cheng et al. 2010]. A characterization of the frequency of each type of page view has also been explored for the case of adult users [Kumar and Tomkins 2010; Cheng et al. 2010]. Nonetheless, none of these studies focus on young users, leaving several research questions open for comparing the browsing behavior of children, teenagers and adults.

2.3.1. Tools to aid children to find information. The development of search technology for children has received increasing attention from the research community in recent

years. A prominent example is the PuppyIR European project[2]. Within the scope of this project, research towards content filtering has been carried out through classifiers that combine topical and non-topical features [Eickhoff et al. 2010]. Graph-based approaches have also been used in this context [Gyllstrom and Moens 2010]. These approaches attempt to filter out Internet websites that are not suitable for children. Query suggestions with topics of interest for children and teenagers derived from social media has been proposed by Duarte Torres et al. [2012]. Language simplification methods based on linear programming optimization have also been proposed to reduce the lexical and grammatical complexity of short texts (in Dutch) [De Belder and Moens 2010]. A framework for interaction with touch interfaces has been proposed by van Dijk et al. [2012]. Showcases applying this research have been shown by Azzopardi et al. [2012a; Azzopardi et al. 2012b; Azzopardi et al. 2012c].

3. METHOD

Search log analysis is one of the main research methods in web search studies to unobtrusively capture the interaction of a large number of users with a search engine [Rieh and Xie 2006]. These analyses have been used extensively to generate statistics of search engine usage, to evaluate web site designs, and to test hypotheses about the effects of different search engine functionality based on user behavioral data [Rieh and Xie 2006; Burton and Walther 2001].

We employed the method *TLA* (transaction log analysis) to conduct web search transaction log analysis [Jansen 2006]. This method was designed for studies aiming at understanding the interactions between searchers, the system and the content provided by these systems. The objective reached through the understanding of the interaction of these three factors may involve improving the design of a search system, improve its accessibility or identifying the users' searching behavior [Jansen 2006]. The interactions refer, within the scope of this method, to a "mechanical expression of underlying information needs" [Jansen 2006] and correspond to the communication between the searcher and the system.

TLA is based on a grounded theory approach in which a systematic discovery of theory is carried out through sampling, comparison, and analysis of data. The resulting models are derived in an inductive manner since the results are grounded in observations of the *real world* instead of being generated by abstract constructs [Jansen 2006]. TLA explores the characteristics of the search sessions and trends are identified from: the queries submitted by the user, from how the queries are modified within the search sessions, from how the result list is explored, and from the type of content that is explored (e.g. multimedia, plain text).

This method consist of three steps:

(1) Collection: Registering and gathering the log data within a predefined time window.
(2) Preparation: The data cleaning process carried out to reduce noise in the data and disregard irrelevant entries for the study.
(3) Analysis: The process of extracting a set of metrics from the data prepared and analysing the metrics in respect to a set of research questions.

We apply these steps for the Yahoo! search logs and the Yahoo! toolbar logs to address Research Questions *RQ-1* to *RQ-3* and *RQ-4* to *RQ-5* respectively. We will describe the application of the three steps in isolation for each dataset.

---

[2]http://puppyir.eu

3.1. Search logs data collection and preparation

The data set employed in this study was extracted from a large sample of the Yahoo! Search logs from May to August of 2010. Only logs from the US market were included. The search interface that users employed correspond to the English Yahoo! US search portal (http://us.yahoo.com), which is meant for all kind of audiences. No cultural differences were considered. In the data collection process, we only included log data from users for whom we could obtain (self-provided) age, gender and US ZIP code. We used the ZIP code in combination with the US census information[3] from the year 2000 in order to annotate users with demographic estimates about their education level (the fraction of the population in a certain age range holding a bachelor's degree or higher). This technique has been previously used in the context of query log analysis by Weber and Castillo [2010; Weber and Jaimes [2011].

The search logs available to the authors consist of rows of event entries, each one associated to a time-stamp. Each event can refer to a user query submission or a click on a web result. In the latter case, the logs also provided the rank position of the clicked result. Each one of these entries was associated to the user profile information, which allows us to group log data from users within the same age group. The age groups are defined in Section 3.1.1.

The data was cleaned by first selecting log entries from users with a valid Yahoo! account. Thus, log entries of users with unspecified birth year, gender and zip code in the Yahoo! profiles were ignored. Log data from users with ill-defined fields were also excluded (e.g. invalid ZIP codes). This filtering step is compulsory in order to be able to identify the age of the user. The resulting data was cleaned further by applying the following criteria:

— Queries containing only a single token and that contain exclusively non-alphanumerical characters
— Queries that were issued by only a single user within a given age group
— Queries containing personally identifiable information, such as credit card numbers or full street addresses

The first criteria was carried out by using a rule-based approach. For instance, regular expressions were designed to detect non-alphanumerical characters in the query string. For the second cleaning criteria we relied on the support of each query, which is obtained by counting the number of users that submit the query within a specific age range.

Only completely anonymous data was used in this study. All the user information that could lead to personal identification in the profiles were inaccessible. Additionally, all references to personal identifiable information (such as credit card numbers, street addresses, persona names) found in the queries were previously removed by Yahoo! and replaced with generic tokens. In this chapter, log entries with any type of such generic tokens (from the data anonymization process) were disregarded, which refers to the third cleaning criteria.

The dataset obtained after applying these cleaning steps for users below 10 years old had a search volume in the order of hundreds of thousands of queries from tens of thousands of users. For users aged 10 or more years old we gathered search volumes in the order of millions of queries from hundreds of thousands of users [4]. As certain aspects of this data set are considered business sensitive, for various metrics we report

---

[3]http://factfinder.census.gov/
[4]Exact statistics about the sizes of the datasets are not reported since they are considered business sensitive information

relative differences between age groups, as opposed to *absolute* differences (i.e. actual click-through-rates on ads).

It is important to mention that a very large fraction of search volume originates from logged in users. Thus, the data analyzed is representative of the population targeted by the search engine used.

3.1.1. Data age segmentation. We aggregated entries from the log data according to the user's birth year. Concretely, we estimated the age of the users by setting the date of birth as the 31 of December of the birth year provided in the user profiles and considering that the search was carried out in 2010. The following age ranges were created:

— early elementary: 6-7 years old
— readers: 8-9 years old
— older children: 10-12 years old
— teenagers: 13-15 years old
— older teenagers: 16-18 years old
— young adults: 19-25 years old
— adults *(i)*: 25-30 years old
— adults *(ii)*: 30-40 years old
— adults *(iii)* : above 40 years old

Our selection of the age groups follows the development changes present in these stages of life. Children from 6 to 7 years old refine their motor skills and start to be involved in social games. Children from 8 to 9 years old start to expand their vision of the world beyond their immediate surroundings. Children from 8 to 12 years old acquire the ability to represent the entities of the world in terms of concepts and abstract representations. Teenagers on the other hand become more interested in social interactions [Kail 2009]. As we will show in this paper, theses stages also have an impact on what children search on the Web, and on the way they interact with the search engine.

Arguably, the user information provided in the Yahoo! profiles is not trustworthy since people can lie about their age, gender or geographical location. Nonetheless, since at last, as early as 2007, Yahoo! has required the consent of a parent or legal responsible for users under 13 years old to create an account [5]. Currently Yahoo! charges a symbolic amount of $.50 to confirm that a guardian is responsible for the child creating the account. Apart from the (small) financial cost, the corresponding time and effort increases the chances of having veridical information for these age groups. Note that even if a small fraction of supposed children users lied about their true age, this is less problematic for general *trends*, though the actual *absolute* numbers will be affected. It is also interesting to notice that in social networks, children tend to lie to make themselves appear older, and this practice is often backed by parents [Richtel and Helft 2011].

3.2. Search logs data analysis

In all of our work, we take a user-centric approach, as we want to provide insights into how young users search online. This means that *all* of our statistics are macro-averages, where metrics are averaged with each user contributing equally, as opposed to micro-averages, where metrics are averaged over all query instances and heavy users will have a bigger importance.

---

[5]http://info.yahoo.com/privacy/us/yahoo/family/details.html

For various parts of our analysis, we make use of the notion of a *search session*. To break sequences of queries and clicks into sessions, we used a simple approach that splits sequences of query and click events into sessions using a sliding window of 30 minutes. A similar approach has been used in several query log studies [Jones and Klinkner 2008; Tyler and Teevan 2010].

For the queries and clicked documents were computed various metrics, which will be explained in the sections where they are analyzed. However, the distinction between navigational and non-navigational queries [Broder 2002] is used in several sections. We describe in this section how this distinction was computed. We combined two different approaches. First, we used the *click entropy* [Weber and Jaimes 2011] to get estimates about how diverse were the clicked results in response to a particular query. Queries that had a sufficient support, a minimum of two occurrences, were judged as navigational if the click entropy was no larger 1.0. This approach works well for head queries (e.g. detects the query "utube" as a navigational query). Additionally, we used a simple heuristic given (query, click) pairs. Note that this heuristic does not label the query as such as navigational, but rather individual (query, click) pairs. Therefore, *facebook* could be non-navigational if the user clicks http://en.wikipedia.org/wiki/Facebook. Our heuristics works as follows: First query and url are tokenized (by whitespaces and dot characters respectively), then tokens are sorted and plurals are stemmed. We label the pair as navigational if the query contains a domain extension (e.g., www, .com, .org), the domain of the url is contained in the query (or viceversa), or the edit distance between the query and the domain of the url is smaller that a threshold value. In the results reported we used two as threshold for queries containing more than four characters). For instance, this method is able to detect the navigational intent of the pair (kids abercrombi, www.abercrombiekids.com/).

3.2.1. Statistical Tests. In our arguments, comparisons between (macro-) averages computed for different groups, say session lengths of children between 6 and 7 and adults between 40 and 70, are core elements. Hence, we were careful to test the various differences we report for statistical significance, using a two-tailed t-test for the equality of means with unequal variance and sample size[6]. We consider a difference to be statistically significant if the probability of the null hypothesis, i.e. the two means being equal, is smaller than 0.1%. It is important to clarify that due to the large volume of data involved in the analysis, most differences were found statistically significant at a 0.1% level according to the two-tailed t-test. The p-values observed for most of the test were smaller than this value, leading to the rejection of the null hypothesis of the statistic under consideration. Nonetheless, we will state in each section the cases in which a given statistic (e.g. macro-average query length) was not statistically significant when comparing age groups. In these cases we will report p-values to provide a clear picture of the results reported. However for simplicity we will omit reporting p-values falling below 0.001.

4. IDENTIFYING AND MEASURING SEARCH DIFFICULTY

Query, click and session characteristics were collected to identify differences in the search process between users of different ages and gender. In the following paragraph we analyze each one of these types of metrics. The focus is on finding metrics that give insight into the *search difficulty* that children face. We will motivate in each section how each one of the metrics explored provide insight into the search difficulties of young users.

---

[6]http://en.wikipedia.org/wiki/Student\%27s_t-test#Unequal_sample_sizes.2C_unequal_variance

Table I: Query length averages by query intent

Age	All		Non-Navigational		Navigational	
	# tokens	#chars.	# tokens	#chars	#tokens	#chars
6 to 7	2.55	16.49	2.80	17.40	2.14	15.16
8 to 9	2.56	16.59	2.77	17.27	2.24	15.91
10 to 12	2.56	16.62	2.81	17.54	2.17	15.49
13 to 15	2.60	16.82	2.84	17.67	2.18	15.67
16 to 18	2.64	17.08	2.86	17.92	2.19	15.68
19 to 25	2.71	17.34	3.03	18.72	2.22	15.16
26 to 30	2.68	17.34	3.02	18.83	2.25	15.32
31 to 40	2.65	17.43	3.00	18.88	2.26	15.76
> 40	2.80	19.05	3.12	20.09	2.43	17.73

4.1. Query length

The formulation of a well-defined query is a crucial part of the search process in IR systems [Downey et al. 2008]. The correlation between query length and IR effectiveness has been explored before [Belkin et al. 2003; Jansen et al. 2000]. On TREC ad-hoc settings, it has been found that longer queries lead to better search performance and user satisfaction [Belkin et al. 2003]. Nonetheless, recent studies show that this result does not always hold on the Web-scale [Downey et al. 2008].

Recently, a strong association between the length of the query and the specificity of the user's query intent has been found [Roul and Sahay 2012; Phan et al. 2007], in which longer queries lead to a more specific and less ambiguous set of results. Thus, the submission of longer queries (in respect to the average length) to the search engine is a strong indicator of the capacity of the user to construct queries that are more specific.

Given previous observations on children's search behavior [Borgman et al. 1995; Fidel et al. 1999], we hypothesized that children have more difficulties to formulate specific queries and that this behavior can lead to ambiguous sets of results, that is, results that are not appropriate for children and that are off topic for their interests. This phenomenon was reported by Bilal [2001; Bilal [2000].

In this work, the obvious two query length metrics were considered: token length and character length. Token length is measured as the number of tokens separated by white spaces, and character length is simply the number of characters (including white-spaces) in the query. Table I summarizes the results obtained by age range. We also discern between the set of navigational and non-navigational queries. Recall that navigational queries are detected based on entropy metrics and (query,click) pair heuristics, as it was explained in Section 3.2. A clear increasing trend was observed from younger to older ages. This result suggests that young users tend to formulate simpler information's goals compare to adult users. Given that the difference margin is larger for non-navigational queries, this result also indicates that young users have difficulties finding the right keywords to formulate more elaborated information needs. This result shows that for ambiguous topics is less likely that young users will retrieve the specific aspects that they are targeting, since the search engine provides content for all type of users, and the volume of information available for adults is significantly larger than the volume offered for young users, particularly children.

With respect to statistical significance, all the macro-averaged statistics were significant using t-test (with p-values $< 1E-5$) except for the following pair: token average for users aged 8 to 9 and users aged 10-12 (p-value = 0.064)

Table II: Fraction of query types

Age	quest.	modals	knowl. quest.	superl.	for kids
6 to 7	2.07%	0.41%	0.16%	0.91%	2.36%
8 to 9	2.56%	0.29%	0.08%	1.48%	1.74%
10 to 12	3.53%	0.58%	0.11%	1.46%	0.97%
13 to 15	3.84%	0.71%	0.16%	1.33%	0.43%
16 to 18	3.33%	0.69%	0.20%	1.15%	0.34%
19 to 25	2.80%	0.49%	0.20%	1.23%	0.32%
26 to 30	2.54%	0.44%	0.16%	1.16%	0.54%
31 to 40	2.19%	0.33%	0.14%	1.09%	0.68%
>40	1.69%	0.24%	0.11%	1.07%	0.31%

4.2. Natural language usage in queries

The aim of analysing the usage of natural language in the queries is twofold: (1) As a mean to retrace children development: Children typically have a greater sense of curiosity [Cecchin 1987], which we hypothesized is reflected in the searches they performed. For instance we expect a greater amount of question queries for users below 10 years old and greater usage of superlative constructs; (2) Children have been observed to pose queries in natural language given their lack of familiarity with the keyword approach of search engines. Greater usage of this type of queries represent evidence of greater difficulty to express complex information needs through keywords which are better suited to modern search engines.

The following query types were created to quantify these phenomena.

(1) Question queries: Queries for which the first token is a question word (how, where, what, ...), or the last character of the query is a question mark (e.g. what is the only immortal animal?)
(2) Modal queries: Queries containing auxiliary verbs as will, won't, don't or modal verbs as shall, should, can, etc. (e.g. I don't want to go school)
(3) Knowledge questions: Queries containing the words *describe*, *about*, *explain*, *define* or *interesting*
(4) Superlatives: Queries containing superlative adjectives (e.g. the fastest dog)
(5) Kids targeted queries: Queries with the terms *for kids* or *for children*

Knowledge queries attempt to measure the fraction of queries intended to fetch a specific explanation about an issue or topic. We decided to explore the usage of superlative queries because this types of construct is commonly employed by children to satisfy the curiosity about certain topic such as in "fastest animal". Note that queries with superlative constructs target in a more concise way a particular aspect of the surrounding world (i.e. objects, situations, persons) compare to other natural language constructs such as adjectives or comparative adjectives [Cecchin 1987]. Superlative queries were detected by looking at tokens with the suffix *est*, and filtering out those matched tokens that are not listed as adjectives in Wordnet[7] or that have a locational meaning (e.g. west).

The kids targeted queries were included because we wanted to explore if children of different ages were using this mechanism to focus the query on content that is suitable for them. Table II summarizes the results obtained by age for the set of nonnavigational queries. On overall, we found that different construct types were preferred at different age ranges. For the case of question queries, we observed that the age ranges 10 to 12, 13 to 15 and 16 to 18 have the highest fraction of question queries,
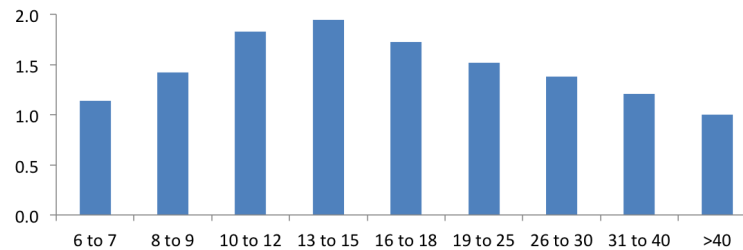
---

[7]http://wordnet.princeton.edu/

Fig. 1: Relative frequency of natural language query types (1 to 4) of each age range against the group of users aged >*40*

the 6 to 7 and 8 to 9 groups do not have a noticeably higher fraction than, say, the 31 to 40 years age range. This result is surprising since in previous studies users below 10 years old have been observed to be more likely to submit this type of queries [Druin et al. 2009b]. We believe this result reflects the higher familiarity of users from 10 to 18 years old with question answering web services like Yahoo! Answers, in which question queries are commonly used. We will elaborate on this finding in Section 5.1.2, in which we address the topic interests of each age range.

The "for kids" part was prominent at the youngest group of users. This result may be an indication that there is self-awareness from this age group of the need of focused content *for children*. Another interpretation of this result is that this age group has greater supervision and aid from parents which are likely to add the *for kids* suffix to the queries. We believe this is an interesting result that requires further research to discern which interpretation is more accurate.

Lack of clear trends was observed for the other categories. Nonetheless, the fraction of superlative queries peaks for children in the 8 to 12 age range. This result was expected since this construct is one of the ways in which children in these ages express their curiosity about the objects that surround them.

We also aggregated the proportion of query types 1 to 4 into a single category for each age group. The ratio between this aggregate and the aggregate found for users above 40 years old was calculated. Figure 1 shows the results. We observed that the greatest usage of natural language is found for teenagers aged 13 to 15, which are 2 times more likely to submit to a search engine this type of queries. High ratios were also observed for users aged 10 to 12 and 16 to 18. We expected to find this behavior mostly on users up to 12 years old, which is the user segment that has been observed to use this type of construct the most [Nahl and Harada 1996; Bilal and Watson 1998]. Nonetheless, we believe this result is also an indication of human development through the search queries. As we mentioned before, teenagers are more prone than other user segments to make use of question answering systems such as Yahoo! Answers, in which a large number of adolescence topics are discussed by the participants (e.g. body changes during adolescence).

When applying the paired t-test we found that the pairs 16-18 / 19-25 (*knowledge question*) and 16-18 / 19-25 (*for kids*) were the only cases in which the results were not statistical significant. The p-values obtained were 0.062 and 0.044 respectively.

4.3. Click position bias

We collect the macro-averaged distribution of the result ranks clicked by the users. The macro rank distribution is computed by estimating the probability of each user to click at each rank position, only taking into account query instances with clicks, and

then averaging the distributions across users belonging to the same age range. We set as cutoff 40 in the estimation of the distributions. That is, we accounted for all the click positions above the 40th ranked position (with better rank than 40). The motivation for this decision is to avoid spam.

Figure 2 presents the distribution of clicks for the first ten results. All the click-through-rates shown are relative to those of users above 40 years old, that is the ratio between the proportion of clicks for the target rank position against the same proportion found for users above 40 years old. A rank of "0" refers to any kind of "special result" which includes links to current news, shopping results or any other high quality content, which is typically only shown for high volume queries. We allowed ourselves to exclude from this plot the relative click-through-rates of users between 19 to 40 years old to make the results displayed from young users (children and teenagers) more readable. On overall, we found that users above 18 years old behaved similarly in terms of click-through rates. On the other hand, the click distribution of teenagers (13 to 15, 16 to 18) and children (6-7, 8-9, 10-12) differed considerably.

Not surprisingly, we found that children users (6 to 12 years old) tended to click on higher ranked results, clicking twice as often as adults on the special rank 0 results. Interestingly, this behavior was even more pronounce for teenagers (13 to 18 years old), for which the ratio was around 2.50. It is important to mention that bias on top ranked results have already been observed in adult users. Hchsttter and Lewandowski [2009] reported that users prefer to click on links that are placed in the first position of the web result list. Craswell et al. [2008] proposed a framework to reduce this bias with log data.

Similarly, for positions 1 and 2 the ratios for children were between 1.10 and 1.30 times as likely as adults to click, while for teenagers the ratios were 1.10 and 1.15 respectively. For lower positions this behavior is reversed, that is, children and teenagers are less likely than adults to click below the third ranked position. On overall, we observed that children are less likely to click below the third rank position than teenagers.

These observations are consistent with previous studies reporting that children have a tendency to explore top ranked results and to avoid exploring thoroughly the list of results when they search for information. Schacter et al. [1998; Fidel et al. [1999] reported this behavior and found that children felt frustrated and overwhelmed by the large amount of data returned by the search systems and particularly by the lack of appropriate content. Our results support these observations, interestingly the same behavior accounted for teenagers. We will show that the frustration is also manifested in other search logs metrics such as click length and session duration.

The statistical tests were carried out based on the macro-average at each rank position. The following pairs were not found statistical significant: 6-7 / 8-9 (at rank 2 with p-value 0.364), 10-12/13-15, 13-15 / 16-19 (at rank 5 with p-value 0.103 and 0.091 respectively) and 8-9 /10-12 (at rank 7 with p-value 0.062). We did not carry out significance tests beyond rank 10.

4.4. Click duration

Previous work showed that a strong signal to detect search success occurs in the form of *long clicks*. Long clicks are defined as those clicks that last 100 seconds of more before another event is registered within the search session. Clicks at the end of a session are ignored from this analysis since they have unknown click duration. We broke down non-final clicks into the three classes suggested by Hassan et al. [2010]: short (0-10 seconds), medium (11-99 seconds) and long ($\geq$ 100 second).

Figure 3 shows that the fraction of long clicks is comparatively low for children of all ages, before it suddenly jumps to a higher level for users in the 19 to 25 age
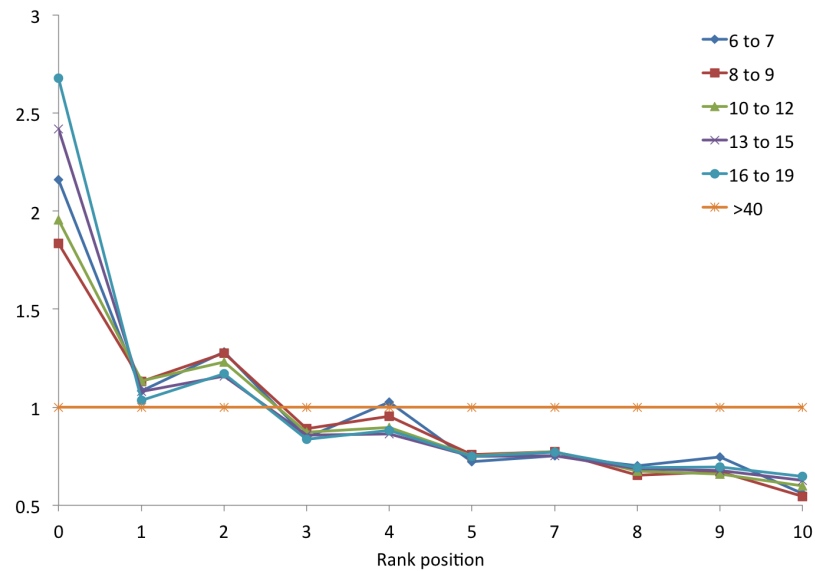
Fig. 2: Relative rank frequency distribution across age ranges. The relative ranks (ratios) are estimated against the age group > 40
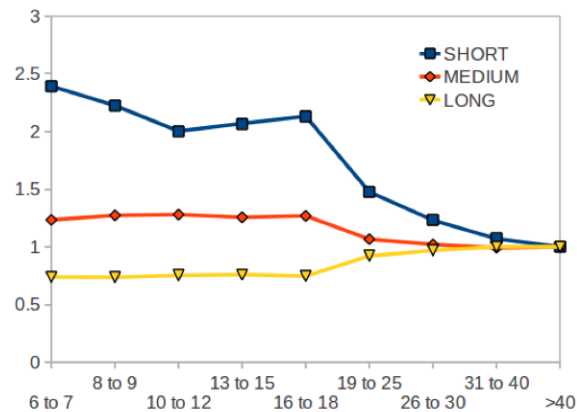

Fig. 3: Distribution of click length across the age groups

range. Frustration of young users during the search process can explain the results observed given that young users tend to abort the clicked pages sooner than adults, thus young users do not spend the same amount of time exploring and parsing the results that are clicked. Frustration may explain this result in the sense that young users may feel overwhelm by the content they retrieve and they stop exploring the websites quickly. This was the case of the two examples derived from the search logs shown in the Introduction.

For the short click macro averages, we found that all pair comparisons were statistical significant (p-values $< 1E - 5$). The following pairs were not statistical significant: 6-7 / 8-9 for medium and long clicks with p-values 0.142 and 0.287 respectively.

Table III: Relative ad click through rates

Age	Click ratio
6 to 7	1.28
8 to 9	1.14
10 to 12	1.04
13 to 15	0.89
16 to 18	0.90
19 to 25	0.84
26 to 30	0.80
31 to 40	1.00
>40	1.20

4.5. Click on ads

We employed the macro-fraction of ad-clicks to quantify how likely it is for a user of a given age range to click on an ad. Since not all the queries trigger advertisements, the estimation was performed only for clicks on results that were generated by queries that had triggered at least one click on an ad. Table III reports the fraction ratios of ad-clicks in respect to the group of adult users between 30 and 40 years old. Values greater than *1* means that users were more likely to click on ads than the age range of adults between 30 and 40 years old.

Surprisingly, we observed higher ratios of ad-clicks for users at very young ages (6 to 12), which suggests disorientation during the search process for these users, since ads are, generally, not targeted at this demographic segment. This observation is in line with previous research that showed that in the context of online games children are also more likely to click on ads, as they fail to recognize them as such [An and Stern 2011; Richtel 2011]. It also reconfirms the findings concerning the position click bias from Section 4.3, given that we only consider ads that are displayed in the top area of the Website. Other advertisements slots are displayed in the bottom and side of the Websites. However, we did not find clear trends in terms on clicks for these two locations.

All age group comparisons were statistical significant (p-values $< 1E - 5$).

4.6. Query assistance usage

Druin et al. [2009a] reported in a detailed case study with 12 participants, that children aged 7 to 12 often ignore the auto-completion and query suggestion facilities provided by search engines. This behavior occurs because of their longer attention on the typing instead of the screen, which make children ignore the queries suggested by the search engine. Figure 4 shows the fraction of queries that were submitted to the search engine as a product of a query suggestion or query correction. Query suggestions are triggered by the search engine when the user is typing the query (e.g. query auto-completion) or as the form of related searches right after the user has submitted the query. The automatic query correction functionality is triggered by spelling mistakes and are commonly displayed by the search engine by informing the user. For instance: *We have included "britney spears" results - Show only "brittnay spears"*.

Figure 4 shows that children were more prone to use query corrections and that they were *not* more likely to make use of query suggestions. This result can be explained by the fact that query corrections are carried out by the search engine automatically, while query suggestions are only displayed and the user need to chose an appropriate query expansion, procedure that has a higher cognitive load. This behavior is inline with Druin et al. [2009a] observations in regard to the the lack of focus experience when children are typing. However, the fact that query suggestions are even less used
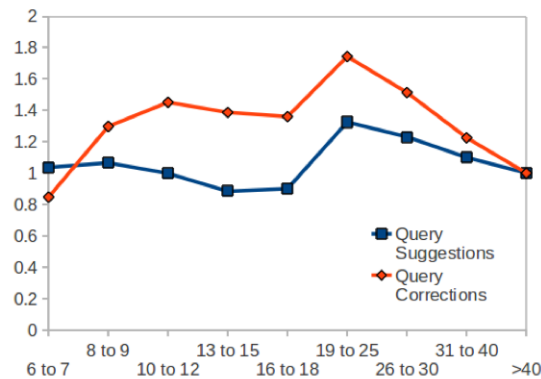
Fig. 4: Query suggestions and correction usage

by teenager users (13 to 18), which have not been observed to have the same focus problems when typing, indicates that the suggestions provided by the search engine are simply not of interest for these users and consequently are not used. This is also the case for users about 40 years old.

We also observed that the younger the users are, the more likely they are to *undo* query suggestions, or in other words, to insist on the incorrect spelling. For instance, by clicking on the option "Show only 'brittnay spears'". The fraction of users aged 6 to 7 to click on such an in-correction was a factor of 1.62 higher than adult users. We believe that this behavior is a consequence of the lack of attention to the screen of young users when they are typing and the click bias. Recall that we found that urls located at the top are more frequently clicked by these users. To undo a query correction is necessary to click on the *undo* link, which is located at the top of the screen. The higher fraction of in-correction observed in young users explains partially why users aged 19 to 25 have a higher usage of this type of query assistance. However, more research is needed in this respect to understand why this particular age group makes more use of query corrections.

4.7. Accidental clicks on explicit content for adults

Children are potentially exposed to adult and explicit material on the Web given its large volume and the lack of parental supervision. Although, we observed lower volume of queries accessing adult explicit content for users below 13 years old (as it will be depicted in Section 5.1) it is important to quantify how often this content is accessed accidentally.

We hypothesized that users clicking by accident on a website with adult content would immediately go back and click on a different web result. Thus, we estimate the likelihood of having a click on a website without adult content after a short click on a website with adult content, in the case when these events are registered during the same search session. Note that this process may occur more that once during the same search session. The last event of the sessions were ignored in the calculations since their click duration are unknown. Adult content was detected using a proprietary classifier based on the Yahoo! Directory[8]. Details about this topic classifier will be given in Section 5.1.

Figure 5 shows the relative frequencies for the event of an accidental short and immediately reverted click on adult content. This figure also shows the relative frequency
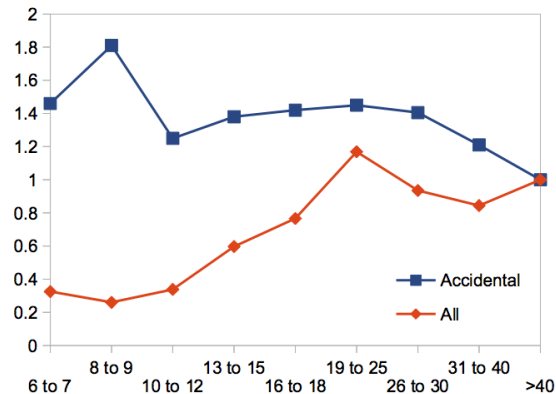
---

[8]http://dir.yahoo.com/

Fig. 5: Relative likelihoods of accidental clicks on adult content websites. The *all* series refer to the relative frequency of clicking on adult content in respect to users above 40 years old

of clicking on adult content in respect to users above 40 years old. On overall, the click on this content increases from the youngest group of users and picks at users from 16 to 19 years old. Note that although children in the 6 to 9 years age range have a comparatively high probability of immediately reverting to a different result, after a (supposedly accidental) click on adult content, their *absolute* probability of clicking on this type of result or of issuing a related query is very low.

The fact that the probability of these accidents-with-immediate correction are higher for children aged 6 to 7 than for children aged 8 to 9 can potentially be explained by the fact that the youngest children might take longer to read an entry page explaining that the site contains adult material and that the visitor needs to be of legal age (typically 18) to view the content. All the age group pairs were statistical significant and all the paired t-test p-values were very small (< 0.001).

4.8. Session characteristics

A sign of search confusion occurs when a user goes back to a query issued earlier in the same session after temporarily exploring different queries. As our sessions were quite short, with an average of 3.51 minutes for users between 8 and 9 years old, it is unlikely that the second occurrence of a query indicates renewal of the earlier information need. More likely, it indicates that the user has not yet fulfilled the earlier information need. We call those queries that are repeated within a session "query refindings", and their fraction is computed as follows. For each user we estimated the fraction of refinding queries inside a session (in respect to the total number of queries inside the session). Then, we averaged these fractions for all the sessions of the same user to generate a per user estimate of query refinding usage. As with all the other metrics reported in this chapter, we reported the macro-average across users. Similarly, a user clicking the same url repeatedly (which can be interspersed with other events) can be seen as an indication that the user is struggling and trying to make up their mind about the most relevant result.

Table IV shows the fraction of refinding queries and clicks. This table also shows two simple measures for the average session length: the length measured in minutes and number of events (queries, clicks and next result page) in a session. It is important to clarify that these estimations excluded sessions containing only one entry (i.e. sessions in which a query was submitted and no clicks were registered).

Table IV: Session characteristics

Age	S.duration	S. length	Query ref.	Click ref.
6 to 7	3.79	3.76	0.24	0.17
8 to 9	3.51	3.71	0.24	0.15
10 to 12	3.63	3.71	0.23	0.14
13 to 15	3.91	3.76	0.26	0.14
16 to 18	4.04	3.82	0.26	0.13
19 to 25	8.20	5.45	0.32	0.22
26 to 30	8.45	5.43	0.30	0.20
31 to 40	8.39	5.28	0.29	0.20
>40	8.42	5.25	0.34	0.24

Table IV shows that the search sessions of children are considerably shorter that for adults. Surprisingly, this statement includes the 16 to 19 age range and the jump to "adulthood" occurs suddenly in the group 20 to 25. This result suggests a greater level of frustration of young users since they tend to quit the search session earlier. This observation is supported by the fact that young users (below 19 years old) have a significant higher proportion of short clicks, as it was pointed out in 4.4. The fraction of query refinding and, in particular, click refinding is *lower* for children. However, rather than to be taken as an indication of a lower level of confusion, this observation is more likely to be due to the fact that children have (considerably) shorter sessions, thus there is simply less opportunity to issue the same query or click.

Non-statistical results (at 0.01%) were found for the following cases:

—Macro average query refinding: 6-7 / 8-9 (p-value 0.231), 13-15 /16-18 (p-values 0.016)
—Macro average click refinding: 10-12 / 13-15 (p-value 0.061)
—Macro average session length: 8-9 / 10-12 (p-value 0.053)

5. TRACING CHILDREN DEVELOPMENT STAGES

Previous work showed that, given enough user search history, attributes such as gender, age and location can be estimated [Jones et al. 2007]. In this work, we looked at a related but different problem: Can we find hints in the query logs that give indications about a child's development stage? (*RQ-3*)

We hypothesize that human development can be traced through the topics and entities targeted by the queries, the gender topic difference, the sentiment expressed in their queries, the reading level of the content clicked and the query vocabulary. In the following sections we explore this hypothesis in detail.

5.1. Topic distribution

We investigated what children search for and how the searches evolve along two dimensions: topics and entities. For the first, we employed a high level classification of topics based on the Yahoo! search directory and a novel fine-grained classification of topics based on the Yahoo! Answers service. For the second, we explored the concrete entities they search and the characteristics of these entities.

5.1.1. High level topic distribution. We used a proprietary classifier to map web pages to entries of the Yahoo! Directory[9]. We used a weighted majority voting scheme on the top 10 organic results returned by the Yahoo! search engine to obtain a classification

---
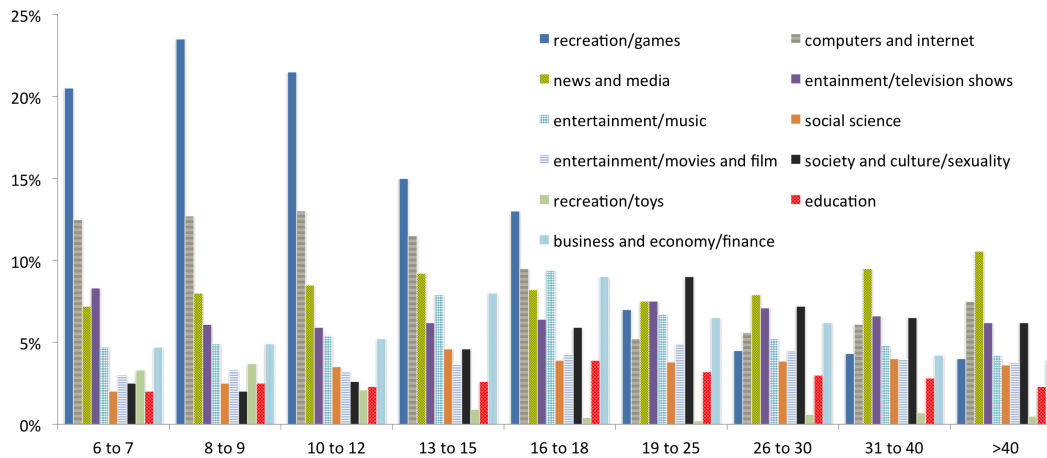
[9]http://dir.yahoo.com/

Fig. 6: Topic progression through the ages

for *queries*, instead of pages. Weber and Jaimes [2011] provides details of this classification method. In total, there were 95 topics. Figure 6 presents the average topic fractions for the 11 most frequent topics searched by users in each age range. Note that we are using queries submitted to the Yahoo! search engine. The Yahoo! Directory is only used to map these queries to categories.

The behavior of Figure 6 is intuitive. Children up to 12 years old had a much higher fraction of queries falling into *recreation/games* than adults and the same holds, at a lower extent, for *recreation/toys*. The interest in music is mostly expressed in the teenager age ranges (13 to 18). The fraction of *business/finance* increases steadily for older users. We also observed a higher diversity of topics for users above 19 years old. In other words, few topics represent a large volume of queries for users below 19 years old, while the queries of older users are more evenly distributed in a larger number of topics.

Even though we are mostly interested in understanding age-related differences, there are also important gender-related differences, even in children [Eccles et al. 1993; Eccles et al. 1990]. We were interested in how gender differences evolve as children grow up. Are gender differences more pronounced in, say, teenagers than in adults? To answer this question we quantified gender differences by looking at the topical distribution for particular age groups. Each one of such topical distribution corresponds to a probability distribution, summing to 100%. We used the 1-norm to quantify the differences between the probability distributions of males and females for each age group. The 1-norm is estimated by summing up the topic proportion values within the topic distribution of each age group[10].

The blue line in Figure 7 shows that the gender differences for children are significantly smaller than for adults. However, many of these gender differences are due to a gender bias introduced by adult content topic. The red line shows the gender differences when this topic has been removed and the remaining topics have been
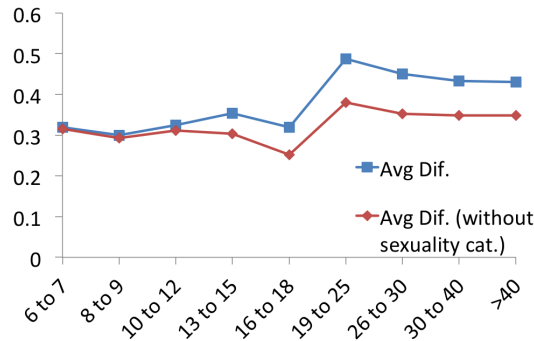
---

[10]http://mathworld.wolfram.com/L1-Norm.html

Fig. 7: Average topic difference between genders through the ages as measured by the $||1||$-norm

renormalized. As can be seen in the plot, this modification removes a large part of the age-related difference between the genders.

The largest differences between the genders were observed in the categories business and economy, computers and Internet and society and *culture/sexuality*. Nonetheless, these differences were significantly higher for males and females above 16 years old, which is the trend that is observed in Figure 7.

Statistical significance was tested by comparing each topic percentage across age ranges (using the paired t-test at 0.1% level of significance). All results were statistically significant with very small p-values (< 0.001).

5.1.2. Fine-grained level topic distribution. In this section, we employed the category structure of the Yahoo! Answers service[11] to have a more detailed classification of the topics targeted by the queries. Concretely, the classification is carried out by submitting the query to the Yahoo! Answers system and the majority vote scheme is used on the categories associated to the top *10* answers for the query. As it was the case with our previous analysis, the Yahoo! Answers system is only employed to classify the queries using the topic hierarchy, recall that the queries studied were originally submitted to the Yahoo! Search engine.

For our analysis, we employed the main categories (e.g. *games & recreation*) and subcategories at depth 1 (e.g. *games & recreation*, *video & online games*). This classification gives us a better overview of the most important topics associated to the concerns and queries that arise in users of different ages. Pinpointing the most frequent topics associated to this type of queries can help designers of search engines for young audiences to focus on certain topic that are critical for these users. For instance, the distribution of topics can be employed to select the most relevant verticals in a search engine according the age of the user.

Concretely, the set of informational queries and a large sample of *how-to* queries were employed in our analysis. *How-to* queries provide a cleaner picture of the concerns that arise in young users. These queries were identified by matching informational queries with the prefix *how-to*. Although more sophisticated mechanisms to identify *how-to* queries have been addressed in previous research [Weber et al. 2012], we believe that this approach provides results with high precision. Given the rapid development of children, we expected to see a rapid change of the topic distribution of the *how-to* queries given that their interests and development stages evolve constantly.

Figure 8(a) depicts the distribution of topics for all informational queries using the global categories. In total, up to 24 general categories and 140 subcategories were
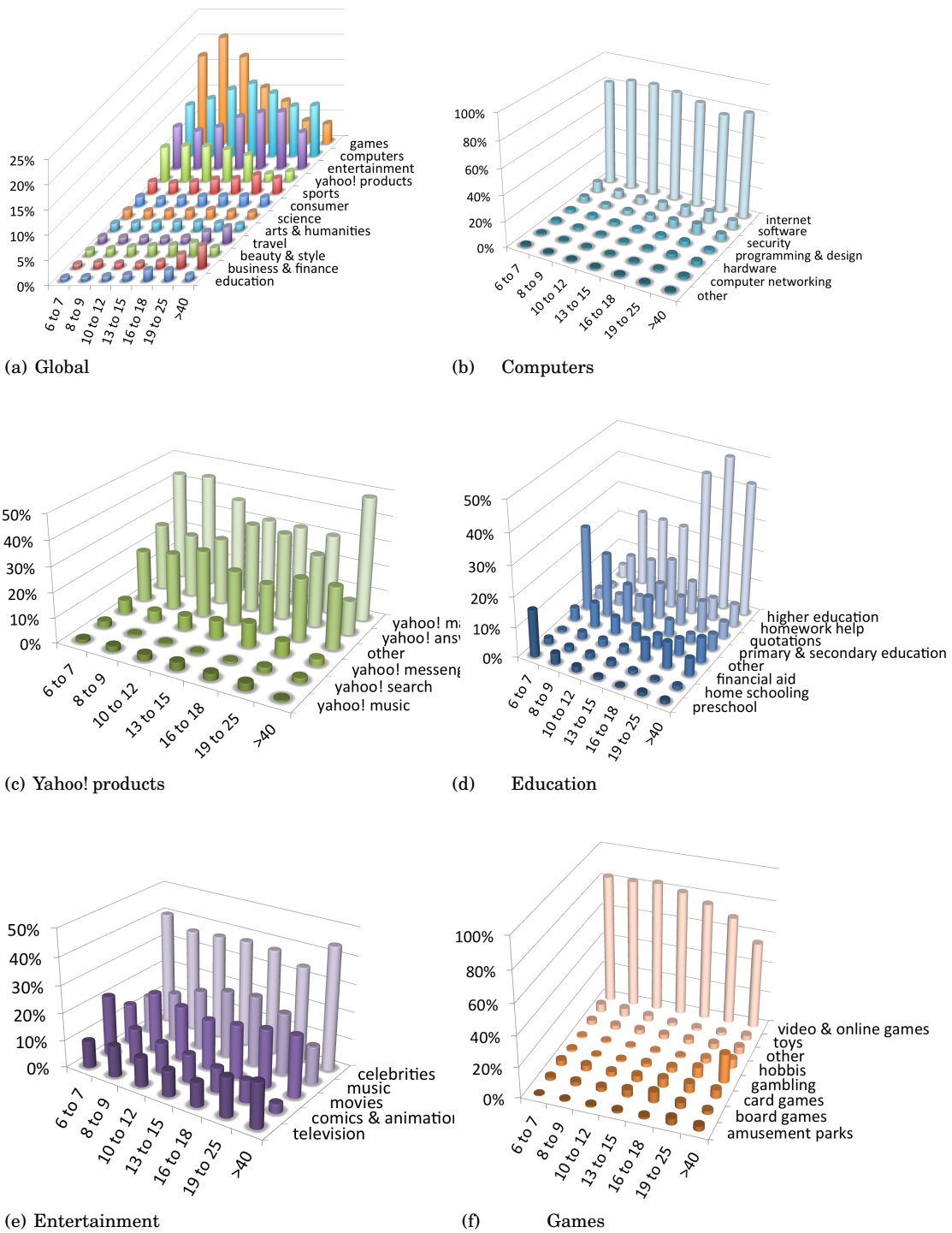
---

[11]http://answers.yahoo.com/

(a) Global

(b)   Computers

(c) Yahoo! products

(d)   Education

(e) Entertainment

(f)   Games

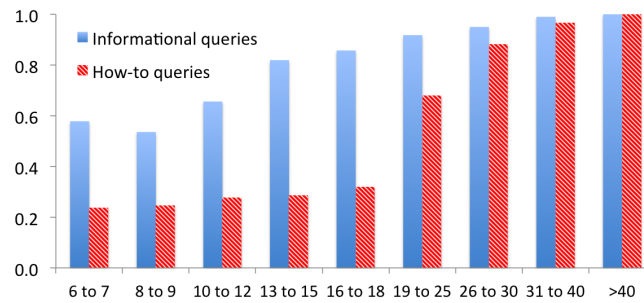**Fig. 8:** Distribution of topics for informational queries

Fig. 9: Pearson's correlation of the topic distribution of each age against the topic distribution of users above 40 years old

identified by mapping the query set to the Yahoo! Answer categories. We observed a strong long tail effect in the topic distribution of queries of adult users. For instance, non-frequent topics (topics that account for less than 5% of the total volume of queries) sum up to 20% of the volume for each one of the groups of users below 12 years old, up to 25% for users up to 19 years old and 38% for users above 19 years old. These observations indicate that the queries of young users are concentrated of a fewer number of topics, while queries submitted by adults are more diverse. This is also reflected in the vocabulary size of young and adult users, as we will show in the next section.

We found that the dominant topics (in increasing order) for users up to 12 years old were: *games & recreation*, *computers & internet*, *entertainment & products*, and *sports*. For teenagers, we observed the same set of dominant topics and *computers & internet*, *entertainment & products*, which had a greater volume of queries than *games & recreation*. This was also the case for users above 19 years old, although the categories *business & finance* and *travel* also gained importance for these users.

We broke down these topics into further detail using the subcategories of the Yahoo! Answers directory. Figures 8(b) to 8(f) depict the results obtained for the subcategories with the highest volume, and the subcategories for which there are significant differences between ages. For *entertainment & products* we found that although the query volume of this topic is large for all the age ranges, the specific *entertainment* subcategory that is searched varies greatly according the age of the user. For instance *comics and animations* is searched 6.5 times more by users aged 6 to 7 than by adults, and up to 2.5 more than teenagers. *Music* is searched 1.9 more by teenagers in respect to children and 1.5 times more in respect to adults. On the other hand, *celebrities* are searched evenly by all the age groups. Figure 8(e) shows the entertainment subtopics distribution for all the age ranges.

For the *yahoo products* category we observed that most of the search volume falls under *email* and *yahoo answers*. In particular, we observed that users up to 12 years old and above 40 years old target 1.5 more frequently queries to access email services than the others age groups. We also observed that queries from users up to 12 years old fall under *Yahoo! Answers* twice as often as adults while teenagers have the greatest percentage of queries for this category (2.6 times the amount of queries than adults submit under this category). This results suggests that the Yahoo! Answers service is a highly valuable resource for young users, particularly teenagers.

For the topic *Computers & Internet* most of the volume was found under *internet*, *software* and *internet security* for the case of adult users. *Internet* is an umbrella for topics related to search engine usage, social networking websites and popular encyclopaedia services as Wikipedia. *Software* is related to the installation, usage, and general support of all kind of computer software. For the exception of *security*, for

Table V: Pearson's correlation of informational query topics and the age of the user

Topic	P. Correlation
yahoo! products	-0.80
games	-0.78
arts	-0.53
pets	-0.40
science & math.	-0.40
computers	-0.15
entertain.	0.09
education & ref.	0.34
beauty & style	0.44
consumer electronics	0.56
sports	0.70
health	0.84
family & rel.	0.86
travel	0.93
business & finance	0.95
news & events	0.96

which the volume of queries for adult users is twice the volume of the other groups, we did not observe clear dissimilarities in the distribution of these subtopics between the age groups. For the case of the games' categories most of the queries were related to *video games* for all the age ranges, as it is illustrated in Figure 8(f).

The results described so far show clear differences and changes in the topic interests of users of different ages. We also quantified objectively the difference between the topic distributions and the difference for each topic between children and adults. The differences between topic distributions were measured using the correlation between the topic distribution of a given age range and the topic distribution of users above 40 years old. The Pearson's correlation coefficient was employed for this purpose. In our context, a positive correlation means that users for the target age range tend to submit queries with the same topics queried by users above 40 years old. On the other hand, lower correlation values mean that the users of the target age range submit queries with different topics that the ones queried by users above 40 years old, which means that is harder to differentiate between both group of users based, solely, on their query topic distribution.

The difference within topics was measured using the Pearson's correlation of each topic and the age of the user. Specifically we estimate the Pearson's correlation for each topic using as variables the lower age in each age group (e.g. 6 for the age group of 6 to 7) and the proportion of the target topic in the given age group. In this case a negative correlation means that the topic is more prone to be used by younger users. A positive correlation means that the topic is more frequently queried by older users. A similar approach has been employed to measure the *trendiness* of words in scientific publications [Hiemstra et al. 2007].

Figure 9 presents the correlation between the topic distributions for the set of informational and how-to queries. This figure shows a marked trend, in which higher correlation values are obtained the older the user is. These observations indicate a clear distinction between the topics queried by young users and adults. Interestingly, we observed that the correlation gap between young and adult age groups is bigger for the how-to queries.

Table V presents the correlation within topics and the age of the user for the set of informational queries. We found that the topics *Yahoo! products*, *games*, *arts* and *science & mathematics* are highly correlated with young ages. On the other hand, topics
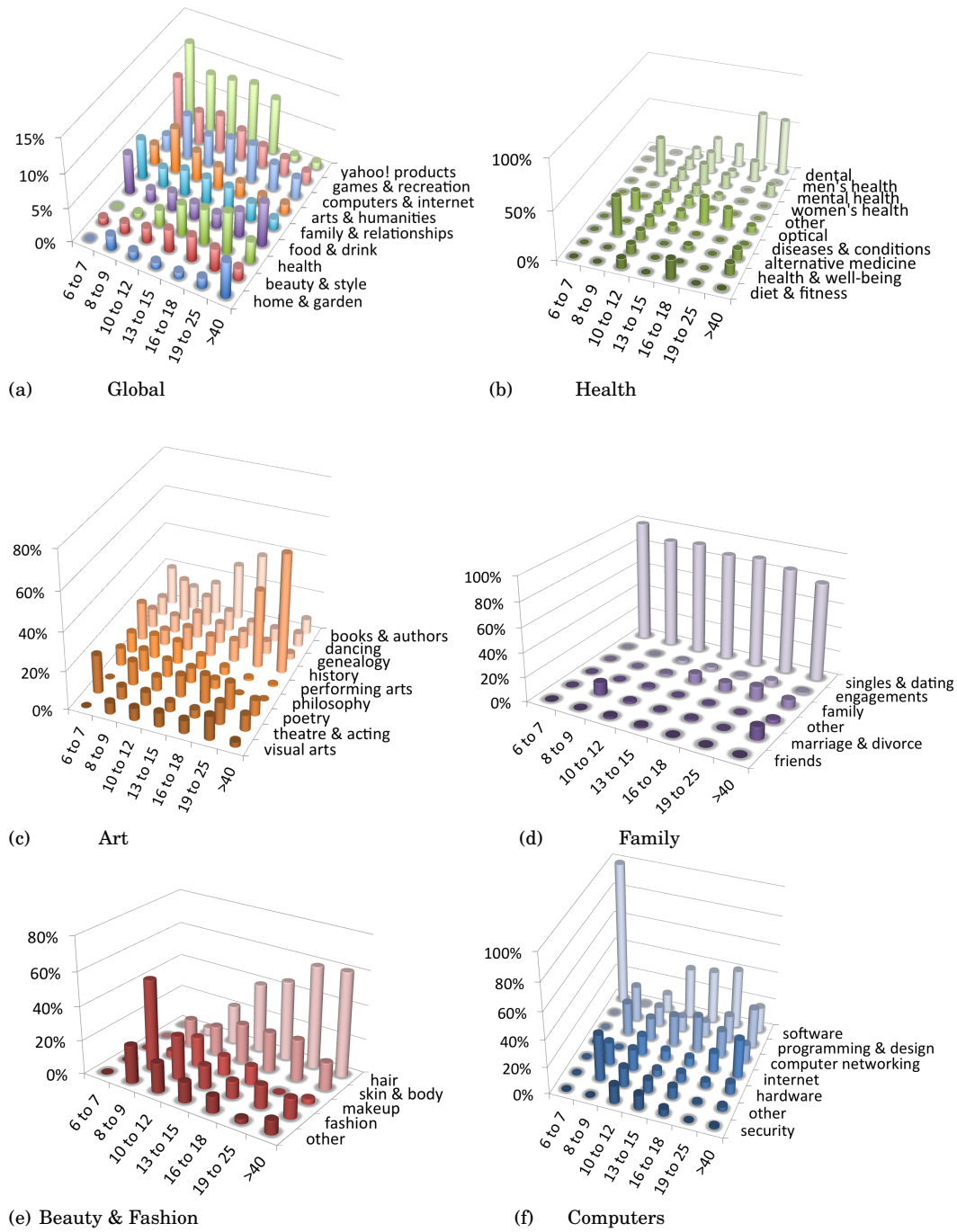
(a)   Global

(b)   Health

(c)   Art

(d)   Family

(e) Beauty & Fashion

(f)   Computers

Fig. 10: Distribution of topics for how to queries

such as *news*, *business & finance* and *travel* are highly correlated with older ages. Categories such as *computers & internet* or *entertainment* did not have a clear correlation with young or old users.

5.1.3. Which specific topics do children target with how-to queries. Figure 10(a) presents the topic distribution obtained with the set of *how-to* queries. As it was the case with the set of informational queries, we found that *yahoo products*, *games* and *computers* represent a large percentage of the volume of queries of young users, with 22%, 14% and 10% of the query volume for users up to 12 years old. The percentages for users between 13 and 18 years old were 18.31% 8.01%, and 11.60% respectively. We also observed a large volume of other topics such as *art & humanities* (10% for users below 10 years old and 9.6% for users between 12 and 18) and *family & relationships* (5.7% for users up to 12 years old and 8% for users between 13 and 18 years old). *Health* and *Beauty* were also prominent for teenagers (10.6% and 7.61% respectively). For adults, the categories *food & drinks* and *home gardening* (12% and 6.7% respectively) had greater importance than for young users.

We also observed the long tail effect difference observed in the previous section between the young and adult group of users. Specifically *non-frequent topics* sum up to 30% of the query volume for users up to 12 years old and around 36% for users above 12 years old. A particular case was the age range of 6 to 7 years old, in which only 8% of the query was associated to *non-frequent topics*.

As we did in the previous section, we explored the subcategories associated to the *how-to* queries. For the subcategories of *family & relationships* we found that most of the queries targeted topics related to *singles & dating*. Surprisingly this was also the case for the youngest group of users (90%, 80% and 78% for users up to 12, 13-18 and above 18 years old respectively). The queries of the youngest group of users classified under this subcategory were manually explored. It was observed that these queries were often regarding early curiosity on dating. Some exemplary queries of this subcategory are: *how to kiss a girl first time*, *how to know what she wants from me? how to find out he likes me?*

Marriage & Divorce was particularly prominent for users above 40 years old (9%). Interestingly, this was also observed for users aged 8 to 9 years old (10%), which are old enough to understand this concept and at the same time being affected by it. For the children group we found that queries falling under this subcategory often reflected concerns regarding parents' relationships (e.g. *how to know if mom work things out with dad*). The subcategory *family* also had a high volume of queries for the case of users above 18 years old (12%).

For the topics related to *art & humanities*, we did not observe a clear dominant subcategory within age groups. In fact, the volume of queries under this category is distributed among most of the subcategories. The distribution of subcategories is shown in Figure 10(c). Nonetheless, we observed large differences among the subcategories across age groups. For instance, queries targeting topics related to the subcategories *books & authors* and *dancing* were up to 3 times more frequent for users up to 12 years old than for adults. The gap between these two sets of users was even greater for the subcategories *performing arts* and surprisingly for *philosophy* (up to 8 times more frequent than adults). Examples of queries under *philosophy* are: *how to take back words? how do you define good and evil*. *History* was more prominent for adults, accounting for 53% of the query volume under this category (4 times more than in the children and teenager ages)

For the category *beauty & fashion* we found that *fashion and accessories* was the dominant topic for users up to 12 years old, accounting on average for 53% of the search volume, which is 3 times the volume found for adults. The dominant subcategory for

Table VI: Pearson's correlation of how-to query topics and the age of the user

Topic	P. Correlation
yahoo! products	-0.82
family & rel.	-0.77
games	-0.75
arts	-0.65
entertain.	-0.52
computers	-0.44
society & culture	-0.21
sports	-0.03
beauty & style	-0.01
consumer electronics	0.00
pets	0.04
health	0.36
food & drink	0.67
business & finance	0.85
home & garden	0.92

teenagers and adults was *hair* with 55% of the search volume (3 times higher that the volume accounted by users up to 12 years old). It was also observed that *makeup* and *skin and body* were prominent for users between 12 and 19 years old (10 times higher the number of searches in respect to adult users). Further details are shown in Figure 10(e).

Finally, for the categories *yahoo! products* and *computers* were found similar trends as the ones reported for the set of non-navigational queries. The results for the *computers* category are depicted in Figure 10(f). As it was the case with the set of non-navigational queries, we measured the correlation with the topics extracted with the how-to queries and the age of the users. Figure VI depicts the results obtained. We found that the topics *yahoo! products*, *games*, *arts* correlate negatively with age. On the other hand, the topics *business & finance*, *health* have a positive correlation with age. This was also the case with the set of non-navigational queries. Interestingly, we found that how-to queries categories such as *family & relationships* or *entertainment* have a negative correlation with age. This result shows that the characteristics of the query (e.g. how-to queries) can influence the topics that correlate best with young users.

5.1.4. Entities targeted by the users' queries. As the topics we used were fairly broad, such as "music" or "finance & investment", we were also interested in obtaining more fine-grained information by looking at the (main) *Wikipedia entity* that a query refers to. To map queries to Wikipedia articles we used the following simple, yet effective, approach: we sent the queries to the Yahoo! search engine and limited the results to results from http://en.wikipedia.org/wiki/. The first result was used as the entity representation for the query. Note that the queries sent to Wikipedia were run fairly recently, though the original queries were submitted about one year earlier. This ensured that even for recent events almost always a Wikipedia page could be found. Table VII shows some examples of this mapping. An overview of the entities searched by young and adults users is presented by the tag clouds in Figure 11 and 12 respectively. These entities correspond only to the non-navigational queries found in our data set. Entities related to adult content were manually removed. One of the advantages of mapping queries to Wikipedia pages is that Wikipedia pages come with a categorical classification and that this classification is both more fine-grained, and in a certain sense orthogonal to our own topic classification (see Section 5.1.1). For example, pages about current celebrities almost always belong to the "Living People" category. Similarly, there are

Fig. 11: Entity tag cloud: 10 to 12 years old

Fig. 12: Entity tag cloud: above 40 years old

Table VII: Examples of queries and their mapped entities

Query	entity
facebook, facebook login	en.wikipedia.org/wiki/Facebook
disney cars games, 2011 cars, cars 2 2011	en.wikipedia.org/wiki/Cars_2
what to do with hummus, ideal protein	en.wikipedia.org/wiki/Hummus
youtuyoutube, youtui, youtuyoutube	en.wikipedia.org/wiki/Youtube
back to school clothes, london school uniforms	en.wikipedia.org/wiki/School_uniform

Table VIII: Entity fractions for Child related content and Living People according to the Wikipedia categories

Age	Children and Kids	Living people
6 to 7	5.81%	8.11%
8 to 9	5.49%	6.97%
10 to 12	3.38%	7.59%
13 to 15	1.46%	9.47%
16 to 18	0.95%	10.86%
19 to 25	0.62%	11.96%
26 to 30	0.63%	11.54%
31 to 40	0.89%	11.09%
>40	0.62%	10.88%

many child-related categories such as "Early childhood education". We used a simple pattern match for the prefixes "child" and "kid" to identify these pages. In Table VIII we present the fraction of entities associated to children content in Wikipedia and the fraction of famous people found in the queries for the age groups.

Given that the celebrities entities often refer to trendy artists, which are more known by teenagers, we expected children and teenagers to have a large fraction of celebrity entities. However, that did not turn out to be the case and the highest fraction of such queries was observed for older user, particularly users in the range 19 to 25 years old. The trend for child-related categories behaved as expected given that the fractions are more pronounced the younger the user is.

All proportions shown in Table VIII were statistical significant when using the paired t-test between pairs age range proportions (all the p-values found were < 0.001).

5.2. Sentiment expressed in queries

Kuhlthau [1999; Kuhlthau [1991] found that uncertainty during the search process, unfamiliarity with technologies, and lack of ownership of the search task (e.g. when children are asked to search a topic by their teacher) can lead to anxiety and frus-

Table IX: Mean query sentiment values scores

Age	Positive	Negative	Diff
6 to 7	1.233	-1.211	0.0216
8 to 9	1.253	-1.237	0.0161
10 to 12	1.257	-1.248	0.009
13 to 15	1.284	-1.274	0.0101
16 to 18	1.274	-1.258	0.0165
19 to 25	1.300	-1.283	0.023
26 to 30	1.302	-1.275	0.026
31 to 40	1.322	-1.297	0.0248
> 40	1.400	-1.376	0.0279

tration, which we hypothesize is manifested in the sentiment expressed in the users' queries. Additionally, children aged 9 to 12 years old have been observed to experience extreme changes of mood [Smith et al. 2011], which we hypothesize may also be reflected in the formulation of queries. Chelaru et al. [2012] also attempted to measure the sentiment of queries by employing a set of queries targeting controversial topics.

To assign numerical scores to the sentiment expressed in the queries, we used the SentiStrength[12] tool developed by Thelwall et al. [2010]. This tool simultaneously assigns both a positive and a negative score to fragments of English text, since users can express both types of sentiments at the same time, such as in "I love you but I also hate you". Positive sentiment strength scores range from $+1$ (somewhat positive) to $+5$ (extremely positive). Similarly, negative sentiment strength scores range from -1 to -5. The tool works by assigning scores to tokens in a dictionary, which includes common emoticons. For example, "love" is mapped to $+3/-1$ and "stink" is mapped to $+1/-3$. Modifier words or symbols can boost the score such that "really love" is mapped to $+4/-1$ (the same for "love!!" or "looove"). The final positive sentiment strength for a bit of text is then computed by taking the maximum score among all individual positive scores. The negative sentiment strength is similarly calculated.

As can be seen in Table IX, sentiment analysis applied to individual queries did *not* reveal the expected trend. It did however reveal that the tendency to use both more positive and negative words in a question *increases* as users get older. This phenomenon is at least partly explained by the fact that they issue longer queries (see Table I) and hence the probability of positive/negative sentiment words appearing is higher. We observed the same behavior when measuring the sentiment of the set of non-navigational queries.

Further research is required to explore our hypothesis in respect to the expression of sentiment through the queries. Current tools to measure text sentiment are not designed for very short pieces of text, which is the case of search queries. The development of more suitable tools to measure sentiment in short texts would be beneficial for this research along the quantification of the sentiment in the content clicked by the users.
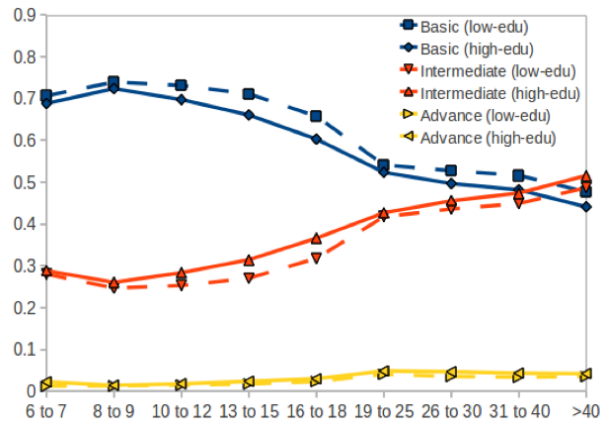
5.3. Reading level of the clicked results

One of the most noticeable factors in children development and its relation to the web search behavior is an improvement in reading skills. As children improve their reading proficiency they will be able to (i) make sense of a wider range of web results, and (ii) potentially understand better the various elements of a web search engine, such as query suggestions or advertisements.

---

[12]http://sentistrength.wlv.ac.uk/

Table X: Examples of websites for each of the three reading levels

reading level	example urls
basic	www.funbrain.com, en.wikipedia.org/wiki/Toy_Story
intermediate	en.wikipedia.org/wiki/John_Wooden, www.sprint.com/
advanced	www.answers.com/topic/mathematics, www.merriam-webster.com

Fig. 13: Reading level across age and average educational level

Collins-Thompson et al. [2011] described a method to personalize web search results based on the readability of the content. They also analyzed the readability of pages accessed by queries targeting content for children, and pages accessed by all type of queries. They found that the former set of queries clearly lead to pages with lower readability complexity. They also found that the higher the readability complexity of the snippets of a web results, the higher is the likelihood of the user to abandon the Website quickly.

To retrace and quantify the improvement in reading level in our data set, we mapped clicked result pages to a 3-scale reading level using Google's "annotate results with reading levels" option.[13] Here, we simply issued the url of the page of interest as a query to Google. In cases where the full url did not return any results, or at least no results with an annotated reading level, we used backtracking by iteratively chopping of parts from the end of the url, hopefully finding a shorter url for which information could be obtained. It is important to clarify that the users studied were unaware of the reading level of the pages clicked, thus no reading level information was displayed to them.

Table X gives a few examples for each of the three reading levels. Note that a single host such as http://en.wikipedia.org can host pages of all three reading levels. We calculated averages across all web pages irrespective of the corresponding query volume: 51.6% of the urls were classified as "basic", 35.8% as "intermediate" and 2.9% as advanced. For 10.2% we could not obtain a reading level with the current approach.

We observed a general and strong trend for the fraction of clicks on "basic" reading level pages, which declines for older users. At the same time, we observed a weak in-

---

[13]http://www.google.com/support/websearch/bin/answer.py?hl=en&answer=1095407

crease for the "advance" level and a strong increase for the "intermediate" level. We also broke down users according to the education level in their self-reported ZIP code, in order to understand which other factors, apart from age, influence the preferred reading level of users. Concretely, we used the census feature "percentage of population of the age of 25 or higher holding a bachelor degree or higher". We sorted users according to this feature and investigated the lowest 20%-tile, and the highest 20%-tile. Figure 13 shows for the fraction of basic reading level pages that children from well educated areas had about 3 years of advantage over children from poorly educated areas. For example, a child from the age range 16 to 18 has a fraction of basic result clicks of 65%. This is slightly lower than the fraction for children in the age range 13 to 15 from well-educated areas, which is 66%, and much higher than the fraction of 60% for other children in the 16 to 18 age range also coming from well-educated areas.

Statistical significance was tested for each proportion (in terms of reading level and educational level) and the proportions were tested in pairs between the ages ranges. All values were statistical significant with p-values < 0.001.

5.4. Query and Click vocabulary

The vocabulary size has been employed in the field of language acquisition as a predictor for reading comprehension and communication [Hellman 2011]. This metric has also been employed to track language development from early childhood to the adulthood [Tonzar et al. 2009]. From the engineering point of view, the vocabulary size has a deep impact on the design of systems for machine translation, speech recognition and part of speech taggers [Church 2011].

In terms of web search, we believe that the vocabulary size can be interpreted as an estimator, in a broad sense, of the capability of users to understand the content of websites, and in the case of web queries as the capability to retrieve content from the Internet through web queries. Thus, a bigger vocabulary size indicates a greater capability to find relevant information. This interpretation is analogous to the capability of carrying out communication being associated to the vocabulary size in the case of natural language communication.

We estimated the vocabulary size by counting the number of distinct words employed in a sample of *10K* queries. To make the comparison fair across age ranges we employed a uniform random sample of the same size for each one of the age groups studied. An analogous procedure was carried out to quantify the *web resource* vocabulary. The *web resource* vocabulary is the average number of distinct urls (or domains) that users of a certain age range access. A larger *domain* and *url* vocabulary is an indicator of a greater diversity of topical interests and greater capabilities for browsing the Internet since a larger number of domains suggests that the user has explored a bigger portion of the Internet.

The results obtained for both query and *web resource* vocabulary size are displayed in Figure 14. Interestingly, we found a clear increasing trend in the vocabulary size of users aged 7 to 25 years old, where the vocabulary stops increasing. Nonetheless, the ratio difference between young children and young adults is less marked than the vocabulary gap reported in natural language. For instance, Moore and Bosch [2009] reported that the average number of words known by a 6 years old person is 14K while for 16 years old is 40K. However, this might be explained by the fact that we are measuring the vocabulary on web queries which are significantly smaller than standard documents. Documents are traditionally employed to estimate vocabulary sizes in natural language.

This result also provides evidence that young users have more difficulties than old users to explore the web given that, on average, they accessed a smaller number of web sites. Note that this is hardly due to a shorter exposure to the Internet since children
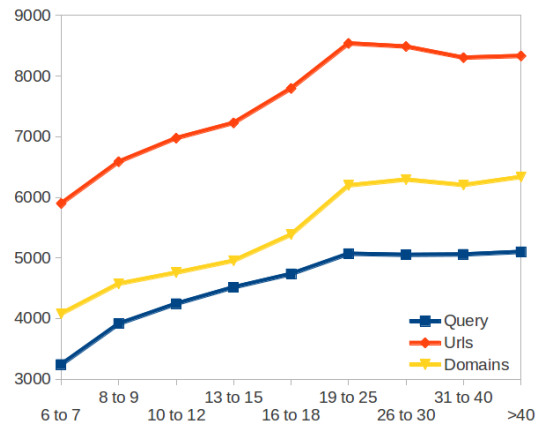
Fig. 14: Vocabulary size across age groups

have been reported to access the Internet from home and even from smartphones. Similarly, the time they spent online has increased dramatically during the last years [14].

This is consistent with our previous findings regarding the click bias of young users. Recall that we found that young users tend to click more often than adults on the very top ranked results, which reduces the potential number of websites they explore. The trends found for the vocabulary sizes can also be explained by the fact that young users are interested in a smaller number of topics, as it was shown in the previous sections (Section 5.1).

From these results, we can conclude that the portion of the Internet that young users explore is significantly smaller than the portion that adults users explore. We believe search engine designers should provide search assistance mechanisms to reduce the query vocabulary gap and to provide new mechanisms to help children improve their accessibility to new content, since they are clicking on a smaller set of links. For instance, web resource recommendation for young users could potentially aid them to explore more content, and especially it can help them to go beyond the simple web result list, which may not be the best approach for serving web content for young audiences.

We tested the differences between the proportions for all the age range pairs for statistical significance. We did not find non-significant results using the paired t-test (all the p-values were < 0.001). We repeated the process for the query, urls and domain vocabulary size, all with similar results.

6. BROWSING BEHAVIOR ANALYSIS

A detailed characterization of the browsing behavior of young users on a large scale has not been addressed previously in the literature. In this section, we describe the Yahoo! toolbar data employed to address research questions *RQ-4* and *RQ-5*. To address these research questions we followed the TLA (Transaction Log Analysis) framework, as it was the case with the Yahoo! Search logs. Recall that this framework is based on the collection, preparation and analysis of log data. We describe each one of the items in the following paragraphs.

---

[14]http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/oct2012/main.pdf

6.1. Toolbar data collection and preparation

We employed logs from the Yahoo! toolbar. The toolbar is a browser application to aid users search and browse the Internet without accessing directly a search engine. The main advantage of analysing the search behavior using logs from the toolbar is that the activity outside the search engine is captured. Traditional search logs only capture the queries and clicks registered during search sessions, on the contrary, the toolbar logs captures searches across different search engines (not exclusively a search engine product) and additional browsing activity carried out before, during and after the search activity.

We collected Yahoo! toolbar log entries for each age range over a time window of four months during 2010. Note that we employed the same time window employed for the Yahoo! Search logs and the same age ranges specified in Section 3.1 of the previous chapter. We collected in the order of tens of thousands of entries for users up to 12 years old, and hundreds of thousands of entries for users above 13 years old. It is important to mention that exact numbers are not reported since they are considered business sensitive information, as it was the case with the search logs.

For the preparation of the data we only employed log entries from users that agreed explicitly to be logged according the usage conditions of the toolbar application[15], and only from users that have a valid Yahoo! account. We employed the definition of valid account described in the previous chapter (Section 3.1). Similarly, queries that could reveal personal information (such as non-frequent names, credit card numbers, telephone numbers) were anonymized prior to the analysis and discarded. It is important to mention that very large volumes of data originates in the Yahoo! Toolbar from logged in users, thus the results presented in this chapter are representative of the underlying population.

Each entry in the toolbar logs has the following structure: {*user ID, timestamp, url, referrer url*}. The *user ID* and *timestamp* were employed to sessionize the data as we did with the search logs. The *referrer url* is the previous web resource accessed by the user and the resource from which the current *url* was found in the entry log. The *referrer url* is used to track consecutive streams of events in the toolbar logs. Contrary to the standard search logs, tracing a linear set of events is not trivial on the toolbar logs given that the user can have several browsing tabs open, and potentially carry unrelated search activity on each one, which are registered in the logs as interleaved events [Huang et al. 2012].

6.2. Yahoo toolbar log analysis

In all the analysis carried out on the toolbar data we created user sessions using a time window of 30 minutes between two consecutive log events. The same strategy has been employed in previous browsing web behavior studies [Cheng et al. 2010; Kumar and Tomkins 2010]. All results presented from the toolbar logs are based on aggregated statistics to preserve the anonymity of the users.

In respect to *RQ-4* we hypothesize that the session characteristics in the toolbar differ by age. Particularly that the amount of search engine usage and Internet browsing changes according the age range of the user. For this purpose, we analyze the sessions of users in the toolbar logs and their characteristics. Specifically, we looked at the length and duration of the sessions and at the proportion of search and browsing activity that is carried out within sessions.

For *RQ-5* we analyze the toolbar sessions in two dimensions: by type of browsing activity and by type of search activity. For the former we employed the taxonomy sug-

---

[15]http://info.yahoo.com/legal/us/yahoo/toolbar/tbeula/tbeula-282.html

gested by Kumar and Tomkins [2010] to classify the possible browsing events that can occur before search activity. For the latter we make the distinction between web search and multimedia search, the latter refers to queries submitted to the Yahoo! Videos or Yahoo! Images search services. Switch patterns of browsing activity to search activity were quantified using the classification of browsing and search activity . These patterns tell us which activities are more likely to occur before search events.

In the following sections we describe in detail the methodology and results obtained for each one of the analysis mentioned above. We will use the following definitions throughout these sections:

Definition 6.1. **Event**: An event refers to a user entering a url in the browser, clicking on a link or entering a query into a search engine. Concretely, it is represented as a tuple {*user ID, timestamp, url, referrer url*}, which is also the representation of an entry in the toolbar log. It is important to mention that redirect links and other log artefacts are not considered events (these entries were properly removed from the data collected).

Definition 6.2. **(Search) query event / (Search) Portal event**: A search event occurs when the *url* matches any of the major search engines (*google, yahoo, bing, aol, ask*) and contains a search query (e.g. *www.google.com/search?q=elmo*). A search portal event occurs when the url matches any of the major search engines but no search query is detected in the url (e.g. *www.google.com*)

Definition 6.3. **(Search) result event**: A search result event corresponds to any *url* clicked from the list of web results obtained for some search query. We detected these events using the *referrer url*. In other words we classify the event as a *search result event* if there exists at least one chain of events connecting the target *url* to a *search query event*.

Definition 6.4. **Browsing event**: A browsing event is any *url* that is not a search query, search Portal, or search result event. We also disregard urls containing queries even if the urls do not match any of the major search engines (e.g. *www.facebook.com*, *www.bbc.com*).

It is important to mention that the results presented in this section were proven to be statistical significant when comparing each one of the children and teenager age ranges against the group of adult users using the two-tailed t-test at a 0.1% level. As it was the case with the search log analysis we report p-values in each section for values that were not proven statistical significant. However, given the size of the data most of the results were statistical significant with very small p-values (< 0.001)

7. SESSION USAGE AND CHARACTERISTICS

The macro-averaged duration in minutes of the sessions and their sizes in terms of number of events were estimated to provide an overview of the browsing behavior of young users on the internet. We also estimate the proportion of search events against browsing events across the age ranges.

Table XI presents the results. We found that the sessions are longer for older users, particularly for users between 19 and 25 years old. On average the sessions of users from 7 to 9 years old contain half the number of events than the average session length of users from 19 to 25 years old. The size of the teenagers' sessions were comparable to the session size of adult users.

Age	Events	Search e.		Browsing e.		Portal e.	
		Volume	Perc.	Volume	Perc.	Volume	Perc.
6-7	31.24	4.67	14.95%	23.55	75.38%	3.02	9.67%
8-9	38.64	4.82	12.48%	30.07	77.82%	3.75	9.70%
10-12	39.36	4.42	11.23%	31.35	79.64%	3.59	9.12%
13-15	55.19	5.29	9.58%	45.78	82.94%	4.12	7.47%
16-19	59.51	5.38	9.04%	50.09	84.18%	4.03	6.78%
19-25	60.16	8.73	14.52%	47.09	78.28%	4.33	7.20%
26-30	56.50	8.28	14.66%	44.35	78.49%	3.87	6.85%
31-40	49.38	7.73	15.65%	37.53	76.01%	4.12	8.34%
>40	41.89	7.29	17.41%	29.89	71.36%	4.71	11.24%

Table XI: Toolbar session and avg. number of search, browsing and portal events per user session and their percentages in respect to the total number of entries
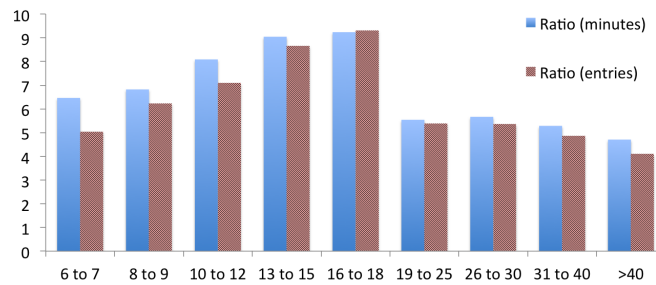

Fig. 15: Ratio of browsing activity against search events in terms of number of events and duration in minutes (textured bars)

The larger sessions of adults is reflected in both greater search and browsing activity. However, we observed that adults have a bigger proportion of search events in respect to browsing events. For instance, for the case of the sessions of users aged 10 to 12, 11% of the events are search events and 80% are browsing events, while for users above 40 years old the proportion of search events is around 17.5% and 71% of browsing events.

Interestingly we also observed a bigger proportion of search activity for the case of users between 8 to 9 and 10 to 12. For teenagers (users between 13 to 15 and 16 to 18), the gap between search and browsing activity is higher: 9% of search events against 84% of browsing events. For the youngest group of users we did not observe large differences in the proportion of search and non-search activity in respect to the sessions of adult users.

Kumar and Tomkins [2010] also explored the proportion of browsing and search activity. Similarly, they reported significant larger amounts of browsing activity in respect to search and portal activity, as we found for most of the age ranges. They reported that on average 9% of the page views correspond to search activity and 21.4% of the page views are derived from search events directly (query submission) or indirectly (browsing a web result). Averages with absolute number of events are not reported. In our findings, the average proportion of search events for users aged 10 to 12, 13 to 15 and above 40 years old were 11.2%, 9.5% and 17.4% respectively.

In general we observed that teenagers are the group of users with the smallest proportion of search. Children and adults had the biggest proportion of search in their sessions. This result is interesting because we were expecting a higher percentage of search activity for the case of teenagers, since these users have greater search capabilities than children.

The lower proportion of search activity found in the children and teenager sessions can be explained by the fact that these users have a strong interest in a few number of topics, as it was shown in the previous chapter in Section 5.1. For instance, gaming is a predominant topic for users below 12 years old. We believe that specialize websites dedicated to this topic can be accessed directly by typing the url in the browser or by bookmarking the website. Recurrent gaming websites such as *Club Penguin, PopTropica and Nick Jr.* are frequently accessed websites by children [16] and it is reasonable to expect that children access these websites directly without searching for them in a search engine.

From the topic preferences observed for teenagers, we believe that a strong bias towards social networking sites (e.g. Facebook, MySpace) lead to the highest proportion of browsing activity, since these web services are easily accessible without the aid of a search engine.

Figure 15 depicts the ratio of browsing activity against search activity in terms of session duration in minutes and session length (number of events). We observed that users from 10 to 12, 13 to 15 and 16 to 18 years old carried out 8 times more browsing activities than search activities. For the case of adults, it was found that users spent 6 times more browsing than searching. For users below 10 years old the ratio was around 5. Interestingly, the session duration ratios were more accentuated than the ratios of the session lengths, which shows that children spent more time exploring the urls.

We also explored differences in the search activity of users by looking at the average proportion of web results and queries submitted within the sessions. We found that a larger number of web results are explored by adults, which is consistent with the findings reported in the previous sections. For instance, the average number of unique queries per session found for users aged 10 to 12 was 0.8, while the number of unique search results (unrepeated search results independently of the query) was 1.5. For users above 19 years old the average found were 1.1 and 3.4, respectively. Note that some values can be lower than 1 because we are accounting for unique queries submitted within the search session and some sessions can contain only instances of the same query or no queries at all. As it was mentioned before, this result indicates that young users explore less the list of web results. This result can also indicates less satisfactory searches for the case of young users since less search results are clicked, given that a similar number of unique queries is submitted within the sessions across all the age ranges.

We tested the statistical significance of all the macro-averages reported in Table XI (number of search events, browsing events and portal events). All the paired tests were found statistical significant with very small p-values (< 0.001). This was also the case for the average duration in minutes, depicted in ratios in Figure 15. The only value that was not statistical significant was the average minute duration for the pair 13-15 and 16-18 (p-value of 0.064)

8. EVENT TO SEARCH QUERY SWITCH PATTERNS

We extracted from the toolbar sessions all the pairs *event → search query* in order to understand the events that are likely to occur before search events. *Event* may refer to a *query event*, *browsing event* or the *start of the session*.

For the *query events* we make the distinction between queries submitted to the standard web search vertical, and queries submitted to the image and video verticals. We will refer to the queries submitted to these two verticals as *multimedia queries*. The motivation for making this distinction is that children have been shown to prefer vi-

---

[16]http://www.ebizmba.com/articles/kids-websites

sual search for certain search tasks [Druin et al. 2010]. We hypothesize that this phenomenon is also reflected in the usage of verticals with visual content.

Cheng et al. [2010] proposed an automatic method to predict search intent based on user browsing activity. They found from the Yahoo! toolbar logs that on average (no age differences were accounted) 19% of the toolbar sessions contained browsing to query search patterns, while 24.3% contains search activity and 75.7% of the sessions contained only browsing activity. We will show that these values are consistent with our findings for the case of adult users, although we observed smaller percentages for the case of young users.

We extended the switch patterns defined by Cheng et al. [2010] to include the usage of multimedia search queries. The definitions are described as follows:

Definition 8.1. *Start of the session → (web/multimedia) query*: This pattern occurs when the search query is the first event of the session.

Definition 8.2. *Web/multimedia query → (web/multimedia) query*: This pattern occurs when the event before the search query is a search query. For the analysis, we only consider the cases in which the previous query is different to the current search query.

Definition 8.3. *Web result event → (web/multimedia) query*: This pattern occurs when a search result event leads to a new search query event. Note that the query of the new search event needs to be different that the query utilized to retrieve the current web result event. We include this pattern because it is reasonable to expect a significant amount of cases in which web results trigger new web searches.

Definition 8.4. *Browsing event → (web/multimedia) query*: This pattern occurs when a browsing event lead to a query search event.

We followed the taxonomy of page views employed by Kumar and Tomkins [2010] to distinguish between the most relevant types of browsing events. Concretely we defined the following browsing event types:

(1) Head listings: events in which the *url* domain is either Amazon, Ebay or Craiglist.
(2) Mailing: events in which the *url* match any of the major email service providers.
(3) Social: View on websites on any of the major social networking websites. We created this list manually by choosing the social networking websites with the greatest traffic volume[17].
(4) Knowledge pages: Websites matching the domain *Wikipedia*.
(5) Multimedia: Websites from Youtube, Hulu, Flicker, Phototobucket,
(6) Other: Websites that could not be classified in any of the previous categories.

8.1. Web search triggers

Figure 16 shows the proportion of the patterns defined in Section 8 for the queries submitted to the Web search vertical. The proportions reported are normalized over the total number of pairs *event → web query* for each age group. We found that a large percentage of browsing activity lead to search events in all the age groups, although this proportion is higher for adults (30% for users below 19, and 34% for users above 25). We also observed that a large proportion of search queries were submitted right at the beginning of the sessions (18.7% for users below 12, and 14% for users above 25). Interestingly, the percentage is higher for children, which suggests that these users start more frequently exploring the internet by using a search engine instead of browsing other resources. This may be due to the fact that the web search engine portal

---

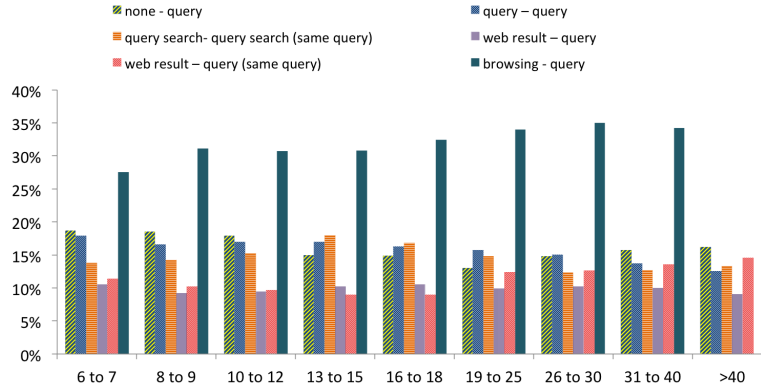[17]http://en.wikipedia.org/wiki/List_of_social_networking_websites

Fig. 16: Proportion of search patterns for all the age ranges

is the default website for most of the commercial browsers. This result also shows that search engines play a bigger role for very young users, since online activities start from the search engine more often than for adult users. We also observed that around 10% of the query events were triggered by exploring web results and 15% of the query events were submitted right after the submission of a different query (e.g. query reformulations).

The following patterns from Figure 16 were not found statistical significant (at 0.1%) when using the paired t-test between age groups: none-query between 6-7/8-9 (p-value 0.537) and 6-7/10-12 (p-value 0.488); browsing-query between 10-12/13-15 (p-value 0.191) and query-query between 10-12/13-15 (p-value 0.033); web result-query (same query) between 6-7/8-9 (p-value 0.412), 10-12/13-15 (p-value 0.288) and 13-15/16-19 (p-value 0.165). All other result pairs were statistical significant at 0.1% with the paired t-test.

Figure 17 presents the likelihood of having a *web query event* given each one of the non-browsing event types. Values reported in Figure 17 were estimated using the fraction between the number of pair events $e \rightarrow web\ search\ query$ over the total number of events e found in the sessions for each age range. For instance, the likelihood of having a *web query event* given a *browsing event* is estimated by dividing the number of pairs *browsing event* $\rightarrow$ *web search query* over the total number of *browsing* events. Consistent with the results reported in Figure 16, children had a higher likelihood than adults of starting a session with web search. The likelihood for children below 12 years old was around 0.55. For teenagers the likelihood was 0.45 and for adults 0.42 (this result was obtained by normalizing over the total number of sessions per user). The probability of web search being spanned by a click on a web result is around 0.04 for users above 25 and around 0.06 for users between 10 to 19 years old. These values indicate that *older* children and teenager users need to refine their searches more frequently than older users since the clicks on web results lead to new searches.

Figure 19 depicts the likelihood found for the browsing patterns. We found that for most of the type of *browsing events* the probability of leading to a web search was relatively low for all the age groups. For instance, the probability of having a search event after a page view on a browsing mail event was around 0.08 for children and adults, while for teenagers the likelihood was around 0.05. For the *social* and *multimedia* browsing events the probability was 0.02 for children and adults, and 0.03 for teenagers. Interestingly, this was not the case for the *knowledge category*, which had
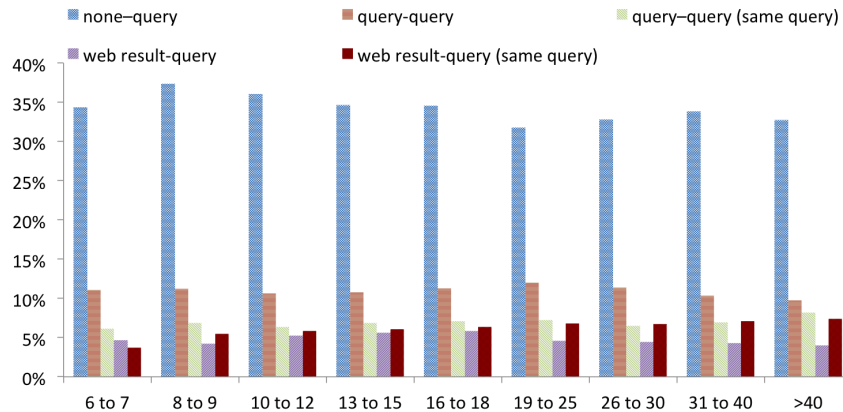
Fig. 17: Search patterns likelihoods (Web search)

a significantly higher chance to occur prior to a web search event. The likelihood was around 0.11 for most of the age groups. For the age group 6-7 the likelihood was 0.16. This result reflects that very young users are keener to carry out web searches after exploring educational content (e.g. Wikipedia) to satisfy their information needs. Even though this result came as a surprise, other studies have reported that Wikipedia is on the top 12 websites more frequently accessed by children aged 5 to 6[18]. These values may also indicate that this group of users requires more explanations or background information to understand the content available in educational resources.

In terms of statistical significance, for Figure 17 we found that the pairs 6-7/8-9 and 10-12/13-15 were not statistical significant for the pattern search query-search query (p-values 0.143, 0.053) and 10-12/13-15 for the search pattern web result - same search query (p-value 0.070).

For the browsing patterns displayed in Figure 19 we found that the pairs 8-9/10-12 were not statistical significant for the knowledge-search query pattern (p-value 0.266) and the patterns involving multimedia and other browsing for the pairs 13-15/16-19 (p-values 0.113 and 0.094).

These observations suggest that providing complementary media related to the topics and content of *knowledge* websites can be highly beneficial for very young user, particularly for improving the readability of educational content. We believe that this aid can be provided in three ways: *(i)* providing definitions to complex words as it has been suggested in domain specific IR [Azzopardi et al. 2012a]; *(ii)* text simplification as it is suggested by Vettori and Mich [2011] and *(iii)* providing related results from different resources to ease the interpretation of the information displayed by using, for instance, results from different genres that may be more familiar and engaging for children (e.g. images, sounds, games).

8.2. Multimedia search triggers

Figure 18 presents the likelihood of having a search event on a multimedia vertical (images [19] or videos [20]) given the non-browsing event types.

On overall, we observed that users between 10 and 19 were 2.4 times more likely to submit queries on a multimedia vertical than adults after most of the non-browsing

---

[18]http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/oct2012/Annex_3.pdf

[19]http://images.search.yahoo.com/

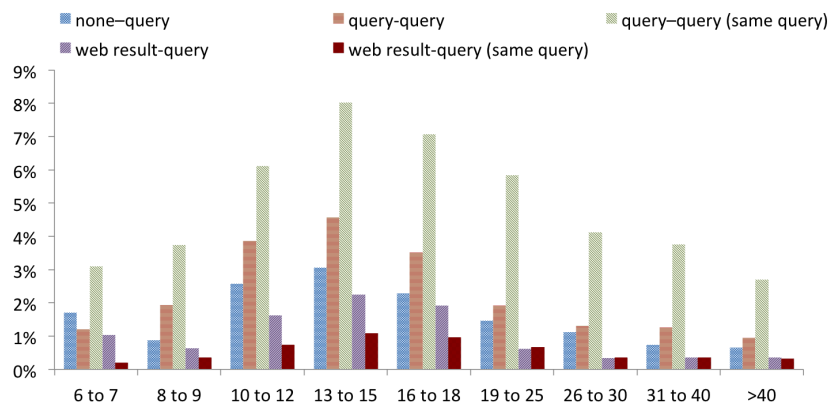[20]http://video.search.yahoo.com/

Fig. 18: Search patterns likelihoods (Multimedia search)
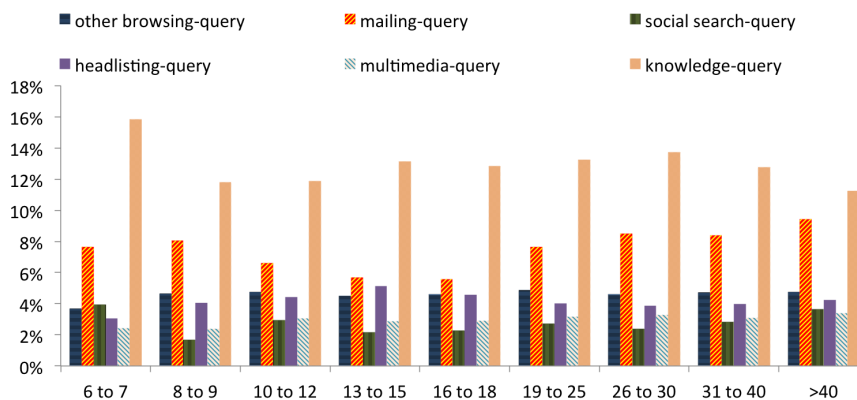

Fig. 19: Browsing pairs likelihoods (Web search)

events. We also observed that the fraction of sessions in which users aged 10 to 19 started the session with a multimedia query was 3% against 0.7% for adults. Even though these percentages are small, the differences in the proportions between age groups are still meaningful, since we are estimating these values from all type of search sessions, in which web (contrary to multimedia) search is predominant.

We observed that the likelihood of performing a multimedia search for users aged 10 years old or below is comparable to the likelihood found for adults. We expected a larger usage of multimedia verticals for the youngest group of users given that the content available on these verticals are generally easier to parse than the websites returned from the standard web search. The low usage found for these users may also be due to the fact that users in these age ranges have been reported to have difficulties to identify the tabs and hyper-links to the non-web verticals [Druin et al. 2010].

We found that *web search* events (from the same query) had the greatest likelihood of leading to a multimedia search query. The fraction of *web query* events that led to multimedia search was 6.9% for users between 10 and 19 and 3.7% for the other age groups. On the other hand, the fraction of *web result* events (from the same query) that led to multimedia search was around 1.0% for users between 10 to 19 and 0.3%
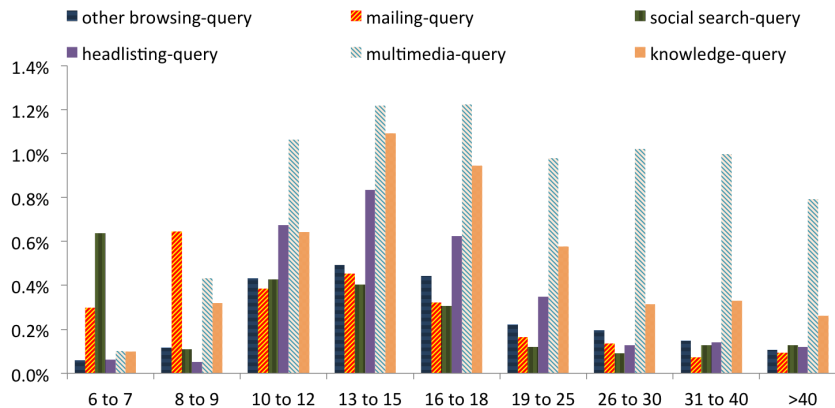
Fig. 20: Browsing pairs likelihoods (Multimedia search)

for the other groups. The higher proportion obtained for the former type of events (*web queries*) can be explained by the fact that several users access multimedia verticals by simply clicking of the tab of image or video search, procedure that automatically send the same query to the multimedia vertical. Nonetheless, as it was mentioned before, these values suggest that children below 10 years old do not use this strategy.

The proportion of web search results that led to multimedia search (with a different query) was also significantly higher on users from 10 to 19 years old. The proportion observed for these users was of 4.5% against 1.3% for the other age groups. This result is interesting because it shows that providing rich media from different genre verticals (e.g. images) can improve the web experience of children in these age ranges (e.g. through aggregated or faceted interfaces).

We observed that the fraction of each one of the browsing events that lead to multimedia search was below 1% for all the age groups, except for the *multimedia browsing* events which classify page views under similar services (e.g. *youtube.com*, *flicker.com*) as browsing activities and account for 1.5% for the young group of users.

Nonetheless, we observed that young users were more likely to perform multimedia search after carrying out browsing activities than adults, as it was the case with the search events. Figure 20 depicts the likelihoods of submitting a query to a multimedia search service after a page view on each one of the browsing types. These results indicate that young users employed more often multimedia search than adults. We consider that improving the accessibility of video and image verticals to young users can highly improve their search experience given that standard web search is one of the main triggers of this type of search. Aggregated search interfaces seem to be an ideal solution to provide results for these users since it eliminates the burden associated to finding the links to the verticals or to locating the vertical results.

All the statistical tests for the proportions reported in Figure 18 and 20 were proven statistical significant with p-values < 0.001. The following patterns in Figure 20 were not found statistical significant: multimedia to search query for the pair 6-7/8-9 (p-value 0.601); headlisting-search query for the pair 6-7/8-9 (p-value 0.548) and mailing-search query for the pair 10-12/13-15 (p-value 0.319).

9. SEARCH TRIGGER CLASSIFICATION

In the previous section we explored the likelihood of a user to carry out web and multimedia search given that the previous event was a browsing or a search event. However,
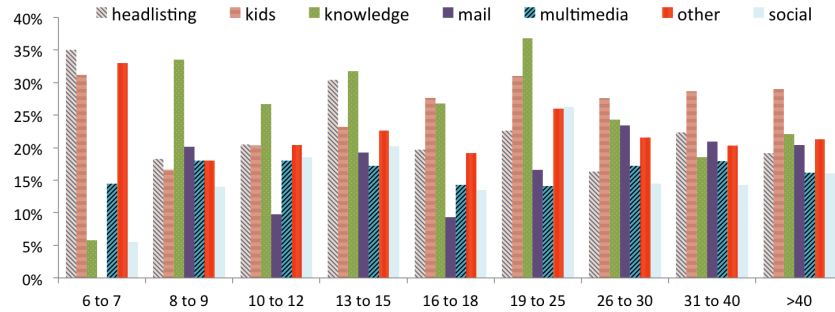
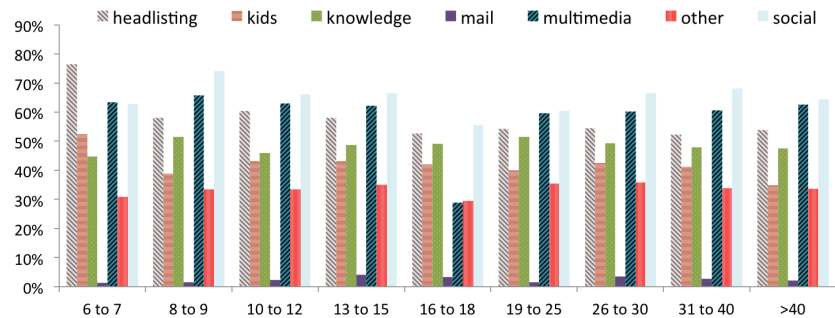**Fig. 21:** Proportion of trigger *1* for the browsing event pairs

Fig. 22: Proportion of triggers *1* and *2* for the browsing event pairs

it is not clear that the browsing or search event triggers the new search. Consider the following two examples to illustrate the cases when a browsing event does and does not trigger a search event: *(i)* a user starts the session by going to a social networking website. The user checks the inbox for new messages. Right after, the user decides to perform a web search to start collecting information for one of their university duties. In this example, the browsing activity (*social browsing*) does not trigger the search event, and the search respond to a change of mood of the user. Now consider the following example: *(ii)* a 9 years old user starts the session by checking the news. The user browses to an article in the science and environment section that discusses new methods to halt rabies. The user does not what rabies is, thus the user decides to copy and paste the word into the search engine box. After understanding that rabies is an illness, the user comes back to the article. In the article the user discovers that rabies is a major issue in certain jungles in South America. The user is intrigued by this place and decides to switch to the search engine once again. This time the user types the query: *animals in south america*. In this example, the two queries submitted are triggered by the information need generated during the browsing activity. In the first case, the query is a highly frequent keyword in the text (i.e. *rabies*). In the second case the query is an information request related to the content of the webpage and it is not explicitly found as a keyword in the page.

We attempt to quantify the proportion of browsing activity that triggers search by classifying the pairs *(event → (search) web query)* into the following categories:

(1) The query is explicitly mentioned in the browsing page.

(2) The query is not mentioned in the browsing page but it is found in the domain of the browsing page.

(3) The query is not found in either the browsing page or its domain.

The first category captures the cases in which the query submitted after the browsing event is explicitly mentioned on the webpage, and it has a high frequency. In the previous example, the submission of the query *rabies* falls under this category since the query is stated in the article several times. The second category captures the cases in which the query submitted target information that is related to the content of the browsing page explored but it not explicitly mentioned on the webpage, which was the case in the previous example for the query *animals in south america*. The third trigger occurs when the user wants to visit a well known resource but the user does not recall the exact url (e.g. *facebook*, *google*).

These categories were proposed by Cheng et al. [2010] after manually inspecting a sample of 200 sessions containing distinct browsing-to-search patterns. A set of technologies were built to detect each case automatically. For the first category, the query keywords are matched with the browsing website and a threshold is set to classify only high frequent keywords. For the second category an analogous procedure is carried out in the domain of the Website.

Figure 21 and Figure 22 depict the proportion of browsing event pairs triggered by the first category (*keyword*) for the set of non-navigational and navigational queries, respectively.

The results reported were obtained by normalizing over the set of event pairs of each browsing type. The proportions were estimated on the set of browsing events that we were able to classify automatically. On overall for each one of the age groups we classified around 40% of the event pairs.

Particularly, we found a high proportion of the *keyword trigger category* for the knowledge browsing event throughout all the age ranges, except for users aged 6 to 7 years old. The proportion for users between 8 to 19 years old was around 0.32, and for older users was 0.23 approximately. The trend observed is consistent with the findings described in the previous section. Recall that the probability of a *knowledge* page view leading to search activity was higher for young users. This finding indicates that a high proportion of these pairs indeed correspond to search triggers and not necessarily to a change of mood or a change of topical interest. However, this trend was not observed for the youngest group. For these users the proportion of search query events triggered by keywords mentioned in the knowledge page was only 5%. This result suggest that these users do not take the same effort shown by older users to carry out follow-up searches related to the content they are browsing.

Interestingly, Figure 21 also shows that the proportion of search triggered by *knowledge search* events decays for users above 25 years old. This may indicate that these users do not need to carry out follow up searches as often as young users given their greater language and topical capabilities. As it was shown in Section 5.1.2 and 5.4 older users have significantly larger query term vocabularies and greater diversity in the topics searched. This result may also be due to the greater search expertise of older users who have better criteria to self regulate their searches and identify when the information collected so far is sufficient or not.

In respect to the *mail* and *multimedia* browsing activities, we did not observe a clear differences between the age groups. For all the age ranges, the proportion of *keyword* search category was around 15% and 18% respectively.

As it was expected, the social browsing event triggered a higher proportion of search activity for the case of teenagers (13 to 19 years old), and the proportion decayed for children up to 12 years old and users above 19 years old. *Headlisting* also triggered a

significant proportion of search activity for all age ranges, particularly for teenagers. This result was expected given that users often carry out follow up searches to find information about items they are planning to purchase (e.g. review, specs pages) or they have interest in.

In terms of statistical significance, all the proportions and pairs in Figure 21 and 22 were statistical significant (once again with very small p-values (< 0.001), except for the following pairs for Figure 21: multimedia for 8-9/10-12 (p-value 0.362), 10-12/13-15 (p-value 0.201) and 8-9/13-15 (p-value 0.045). For Figure 22: knowledge for 13-15/16-18 (p-value 0.062); social for 13-15/16-18 (p-value 0.304); multimedia for 10-12/13-15 (p-value 0.421); kids for 10-12/13-15 (p-value 0.415), 13-15/16-18 (p-value 0.203) and 10-12/16-18 (p-value 0.354) were not statistically significant.

10. CONCLUSIONS

10.1. Findings summary

In respect to the research question *RQ-1* we observed clear evidence of search difficulty in young users. The shorter average query length observed in young users along with the greater usage of natural language shows that these users have difficulties formulating specific queries with keywords, which is the main mechanism to query in state-of-the-art search engines. The larger proportion of short clicks observed is also an indication of search difficulty since it suggests that pages are abandoned quickly. In terms of click behavior the position click bias for children is worth pointing out. This bias also leads to a higher fraction of ad clicks and to a higher fraction of cases where (useful and correct) spelling corrections are undone by the user when clicking on, say, "Show only '*brittnay spears*' ". Both of these indicate that very young users have a tendency to "click whatever is presented at a prominent position" which has implications for the design of an appropriate search interface. This behavior, along with the trust young users have on the veracity and quality on the content published on the Internet, contributes to clicks on content that is not suitable for young users. This was observed as a greater proportion of accidental clicks on adult content and a larger amount of advertisement clicks which potentially lead to more search frustration.

In respect to *RQ-2* we found notable difference between children and adults in terms of search and browsing behavior. However, it was surprising to find small differences between children and teenager users. For instance, users between 16 and 18 behaved more like children between 8 to 9 than young adults in the 19 to 25 age range. This "sudden jump to adulthood", albeit not for all features, could potentially be explained by children leaving home and starting college or a job. Some features in which children and teenagers differ by a large margin were: topic distribution, query vocabulary size and the proportion of question queries submitted to the search engine.

For *RQ-3* we observed that several aspects of children and teenagers development were reflected in the search logs. For instance, we found a direct correlation between the topic distribution and age. Interestingly, we observed a wide variety of topics in the distribution of topics for adults users while for children and teenagers the queries concentrated on a small set of topics. Stereotypical topics such as gaming and entertainment were observed in children and teenagers respectively. However, we observed an unexpected higher interest in emailing and other Yahoo! products such as Yahoo! Answers for the case of children. The latter was also prominent for the teenager segment of users.

The development of children was also reflected in clear differences in the query vocabulary size and the readability level of the content accessed. Features that we hypothesized could reflect human development, such as the sentiment express in queries,

were not conclusive and more research is required in this regard and other effectual aspects of search.

In regard to *RQ-4* a larger proportion of browsing activity compared to search activity was observed for all age groups, however we found that younger users have a greater proportion of browsing events. 85%, 80% and 81% of the toolbar session entries were browsing activities for children, teenager and adults respectively. We believe this result is due to the topic bias towards gaming and social services of children and teenagers, services that are easily accessible without a search engine. This result also shows that current search engines need to address the search difficulties that children and teenagers face on the Internet to attract more this segment of users.

For *RQ-3.2* we found that most of the users engaged search from the beginning of the session, although this behavior was more pronounced in the youngest group of users. This was also the case for teenagers in the case of multimedia search. We found that knowledge intense websites (e.g Wikipedia) had the highest likelihood of triggering search for all the age ranges among all the browsing activities. Interestingly, teenagers were found to be twice as likely to submit queries on multimedia services than other age groups. Web search was the most likely activity to trigger multimedia search. For the browsing activities knowledge and headlisting were the two type of pageviews with the largest likelihood of triggering multimedia search.

In the following section we describe the implications of these results on the development of IR tools for children and we point out directions for future research.

10.2. Recommendation for the development of information tools for children

A call for better query assistance functionality has been pointed out in previous research [Bilal 2002; Druin et al. 2009a]. In our work, we also found evidence of the need of this functionality. From our observations, young users (particularly up to 12 years old) would greatly benefit for query expansion and query recommendation mechanisms, since their average query length is the smallest from all the age groups. They also had the smallest query vocabulary size. Query suggestions with information aspects (i.e. query senses) tailored with topics of interest for children would greatly help children to narrow down their searches, and it would improve their chances to find information that is on topic. These tools would also improve the chances of returning results that are better suited for them. We expect that high quality suggestions would greatly increase the usage of these tools by children. Note that the query strategies preferred for each age group should also be considered in the design of digital libraries aimed at school children.

For the case of teenagers we found interesting the large usage of question queries, which suggests the need of providing robust mechanism to parse this type of queries in order to understand in deep their information needs. Alternatively, resource selection mechanisms to identify when it is appropriate to return results from specialized question answering systems based on the topic of the query are also relevant, especially since we observed that children and teenagers make extensive use of these services (e.g. Yahoo! Answers).

The main topics identified in the topic distribution for each age range are valuable for the design of aggregated search interfaces, since the more prominent topics can be associated to specific verticals (e.g. games, school, entertainment). In this regard, resource selection methods and novel aggregation paradigms for young users are worth to be explored.

Aggregated interfaces are a promising alternative to present results from multimedia verticals related to the web results. We believe that this type of interfaces can greatly benefit the search experience of children and teenagers. Recall that we observed a large likelihood of query submission to these services after carrying out

browsing activities. Tools to find complementary material for educational content (e.g. Wikipedia, Simple Wikipedia) can also support the information collected by young users. These tools may involve finding related audio-visual material and tools for language simplification.

Teachers and experts of child care can also benefit from the findings of this study. For instance, we observed that children seem to have difficulties to recognize advertisements on the web. Teachers can help children in the classrooms to improve the awareness towards this type of content. Additionally, the topics and media genres preferences for each age range can be utilized in the design of specialized educational literacy for these users.

From our findings, methods to improve the access to knowledge search services (with content readable by children) and multimedia content are urgently required by the youngest group of users.

10.3. Recommendations for future research

Even though we collected clear evidence to address the research questions posed in this chapter, several open questions still require further research. We characterize the search behavior of users based on their age, however it is unclear how this behavior varies based on the topic that is searched and the specific type of information need (e.g. open and close informational needs). A natural follow up of this work would include a search analysis conditioned not only on age but also on topic.

We explored the proportion of clicks on content of different reading difficulty by considering the age and the average educational level in which the user is located, by combining the search logs and the U.S census data from 2000. The exploration of the search difficulties and search behavior differences of users within other demographic features is a highly relevant research line that can provide clues on how to improve the search experience of specific niche of users. For instance, knowing the specific search difficulties of users per educational level, region and income would lead to the possibility of displaying, if the user wants to, specialized results that address specific search difficulties. Note that we are not suggesting to blindly show users specific results according their demographics, instead we suggest to provide the option to retrieve these specialized results to all the public.

Demographic features such as race and language proficiency require special attention to provide adequate tools aimed at improving the content readability. Cultural differences in terms of search behavior, and particularly cultural topic preferences represent another important, yet not well-studied, area in information retrieval.

The impact of aesthetic characteristics of the websites visited by young users is yet to be explored, particularly on the large scale and its relation to the problems that arises during the search process of young users. For instance, page layout, amount of multimedia content, number of images, font types, font sizes, text and background color are all features that need to be explored in future studies. Measuring the impact of these design parameters would lead to a better understanding in how to design content that is more engaging for children.

More research is also needed to understand the affective factors that arise during the search process for the case of very young users. We quantify the amount of sentiment in the queries, however no trends were found. Quantifying the sentiment in the content accessed by children and exploring its correlation to the click duration can provide a better understanding in this respect.

As it was pointed out in the previous section, research towards mechanisms to automatically select search services relevant to the topic based on the query and age of the user would be highly beneficial. New paradigms to aggregate content from these ser-

vices are also needed to simplify the exploration of results and the engagement during the search process.

ACKNOWLEDGMENTS

REFERENCES

AN, S. AND STERN, S. 2011. Mitigating the effects of advergames on children; do advertising breaks work? *Journal of Advertising 40,* 1, 43 – 56.

AZZOPARDI, L., DOWIE, D., DUARTE, S., EICKHOFF, C., GLASSEY, R., GYLLSTROM, K., HIEMSTRA, D., DE JONG, F., KRUISINGA, F., MARSHALL, K., MOENS, S., POLAJNAR, T., VAN DER SLUIS, F., AND DE VRIES, A. 2012a. Emse: supporting children's information needs within a hospital environment. In *ECIR'12*. Springer-Verlag, Berlin, Heidelberg, 578–580.

AZZOPARDI, L., DOWIE, D., AND MARSHALL, K. A. 2012b. Yoosee: a video browsing application for young children. In *SIGIR '12*. ACM, New York, NY, USA, 1017–1017.

AZZOPARDI, L., DOWIE, D., MARSHALL, K. A., AND GLASSEY, R. 2012c. Mase: create your own mash-up search interface. In *SIGIR '12*. ACM, New York, NY, USA, 1008–1008.

BEHESHTI, J., LARGE, A., AND TAM, M. 2010. Transaction logs and search patterns on a children's portal/journaux de transaction et modes de recherche sur un portail web destiné aux enfants. *Canadian Journal of Information and Library Science 34,* 4, 391–402.

BELKIN, N. J., COOL, C., KELLY, D., KIM, G., Y. KIM, J., J. LEE, H., MURESAN, G., C. TANG, M., AND J. YUAN, X. 2003. Query length in interactive information retrieval. In *SIGIR '03*. ACM, New York, NY, USA, 205–212.

BILAL, D. 2000. Children's use of the yahooligans! web search engine: I. cognitive, physical, and affective behaviors on fact-based search tasks. *JASIS 51,* 7, 646–665.

BILAL, D. 2001. Children's use of the yahooligans! web search engine ii. cognitive and physical behaviors on research tasks. *J. Am. Soc. Inf. Sci. Technol. 52,* 2, 118–136.

BILAL, D. 2002. Children's use of the yahooligans! web search engine iii. cognitive and physical behaviors on fully self-generated search tasks. *J. Am. Soc. Inf. Sci. Technol. 53,* 13, 1170–1183.

BILAL, D. AND WATSON, J. S. 1998. Children's paperless projects: Inspiring research via the web. In *IFLA. Conference*.

BORGMAN, C. L., HIRSH, S. G., WALTER, V. A., AND GALLAGHER, A. L. 1995. Children's searching behavior on browsing and keyword online catalogs: the science library catalog project. *JASIS 46,* 9, 663–684.

BOWLER, L., LARGE, A., AND REJSKIND, G. 2001. Primary school students, information literacy and the web. *Education for Information 19,* 3, 201–224.

BROCH, E. 2000. Children's search engines from an information search process perspective. *School Library Media Research 3*.

BRODER, A. 2002. A taxonomy of web search. *SIGIR Forum 36,* 3–10.

BURTON, M. C. AND WALTHER, J. B. 2001. The value of web log data in use-based design and testing. *Journal of Computer-Mediated Communication 6,* 3, 0–0.

CECCHIN, G. 1987. Hypothesizing, circularity, and neutrality revisited: An invitation to curiosity. *Family process 26,* 4, 405–413.

CHELARU, S., ALTINGOVDE, I. S., AND SIERSDORFER, S. 2012. Analyzing the polarity of opinionated queries. In *ECIR'12*. Springer-Verlag, Berlin, Heidelberg, 463–467.

CHENG, Z., GAO, B., AND LIU, T.-Y. 2010. Actively predicting diverse search intent from user browsing behaviors. In *WWW '10*. ACM, New York, NY, USA, 221–230.

CHILD TRENDS DATA BANK. 2013. Home computer access and Internet use.

CHURCH, K. 2011. How many multiword expressions do people know? In *Proceedings of the Workshop on Multiword Expressions: from Parsing and Generation to the Real World*. MWE '11. Association for Computational Linguistics, Stroudsburg, PA, USA, 137–144.

COLLINS-THOMPSON, K., BENNETT, P. N., WHITE, R. W., DE LA CHICA, S., AND SONTAG, D. 2011. Personalizing web search results by reading level. In *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 403–412.

CRASWELL, N., ZOETER, O., TAYLOR, M., AND RAMSEY, B. 2008. An experimental comparison of click position-bias models. In *WSDM '08*. ACM, New York, NY, USA, 87–94.

DE BELDER, J. AND MOENS, M.-F. 2010. Integer linear programming for dutch sentence compression. In *Proceedings of the 11th international conference on Computational Linguistics and Intelligent Text Processing*. CICLing'10. Springer-Verlag, Berlin, Heidelberg, 711–723.

DOWNEY, D., DUMAIS, S., LIEBLING, D., AND HORVITZ, E. 2008. Understanding the relationship between searchers' queries and information goals. In *CIKM '08*. ACM, New York, NY, USA, 449–458.

DRUIN, A., FOSS, E., HATLEY, L., GOLUB, E., GUHA, M. L., FAILS, J., AND HUTCHINSON, H. 2009a. How children search the internet with keyword interfaces. In *IDC '09: Proceedings of the 8th International Conference on Interaction Design and Children*. ACM, New York, NY, USA, 89–96.

DRUIN, A., FOSS, E., HATLEY, L., GOLUB, E., GUHA, M. L., FAILS, J., AND HUTCHINSON, H. 2009b. How children search the internet with keyword interfaces. In *IDC '09: Proceedings of the 8th International Conference on Interaction Design and Children*. ACM, New York, NY, USA, 89–96.

DRUIN, A., FOSS, E., HUTCHINSON, H., GOLUB, E., AND HATLEY, L. 2010. Children's roles using keyword search interfaces at home. In *Proceedings of the 28th international conference on Human factors in computing systems*. CHI '10. ACM, New York, NY, USA, 413–422.

DUARTE TORRES, S., HIEMSTRA, D., AND SERDYUKOV, P. 2010. An analysis of queries intended to search information for children. In *Proceedings of the third symposium on Information interaction in context*. IIiX '10. ACM, New York, NY, USA, 235–244.

DUARTE TORRES, S., HIEMSTRA, D., WEBER, I., AND SERDYUKOV, P. 2012. Query recommendation for children. In *CIKM '12*. ACM, New York, NY, USA, 2010–2014.

DUARTE TORRES, S. AND WEBER, I. 2011. What and how children search on the web. In *CIKM '11*. ACM, New York, NY, USA, 393–402.

ECCLES, J., WIGFIELD, A., HAROLD, R. D., AND BLUMENFELD, P. 1993. Age and gender differences in children's self- and task perceptions during elementary school. *Child Development 64,* 3, 830–847.

ECCLES, J. S., JACOBS, J. E., AND HAROLD, R. D. 1990. Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues 46,* 2, 183–201.

EICKHOFF, C., SERDYUKOV, P., AND DE VRIES, A. P. 2010. Web page classification on child suitability. In *CIKM '10*. 1425–1428.

FIDEL, R., DAVIES, R. K., DOUGLASS, M. H., HOLDER, J. K., HOPKINS, C. J., KUSHNER, E. J., MIYAGISHIMA, B. K., AND TONEY, C. D. 1999. A visit to the information mall: Web searching behavior of high school students. *Journal of the American Society for Information Science 50,* 1, 24–37.

GOEL, S., HOFMAN, J. M., AND SIRER, M. I. 2012. Who does what on the web: A large-scale study of browsing behavior. In *ICWSM*. AAAI, Toronto.

GOSSEN, T., LOW, T., AND NÜRNBERGER, A. 2011. What are the real differences of children's and adults' web search. In *SIGIR '11*. ACM, New York, NY, USA, 1115–1116.

GUNN, H. AND HEPBURN, G. 2003. Seeking information for school purposes on the internet. *Canadian Journal of Learning and Technology/La revue canadienne de lapprentissage et de la technologie 29,* 1.

GYLLSTROM, K. AND MOENS, M.-F. 2010. Wisdom of the ages: toward delivering the children's web with the link-based agerank algorithm. In *CIKM '10*. ACM, New York, NY, USA, 159–168.

HASSAN, A., JONES, R., AND KLINKNER, K. L. 2010. Beyond DCG: user behavior as a predictor of a successful search. In *WSDM '10*. ACM, New York, NY, USA, 221–230.

HELLMAN, A. B. 2011. Vocabulary size and depth of word knowledge in adult-onset second language acquisition. *International Journal of Applied Linguistics 21,* 2, 162–182.

HIEMSTRA, D., HAUFF, C., DE JONG, F., AND KRAAIJ, W. 2007. Sigir's 30th anniversary: an analysis of trends in ir research and the topology of its community. *SIGIR Forum 41,* 2, 18–24.

HUANG, J., LIN, T., AND WHITE, R. W. 2012. No search result left behind: branching behavior with browser tabs. In *WSDM '12*. ACM, New York, NY, USA, 203–212.

HCHSTTTER, N. AND LEWANDOWSKI, D. 2009. What users see structures in search engine results pages. *Information Sciences 179,* 12, 1796 – 1812. ¡ce:title¿Special Section: Web Search¡/ce:title¿.

JANSEN, B., BERNARD, M., JANSEN, J., SPINK, A., AND SARACEVIC, T. 2000. Real life, real users, and real needs: A study and analysis of user queries on the web. *Information Processing & Management 36*, 207–227.

JANSEN, B. J. 2006. Search log analysis: What it is, what's been done, how to do it. *Library & information science research 28,* 3, 407–432.

JOCHMANN-MANNAK, H., HUIBERS, T., LENTZ, L., AND SANDERS, T. 2010. Children searching informa-
    tion on the internet: Performance on children's interfaces compared to google. In *Towards Accessible
    Search Systems*, P. Serdyukov, D. Hiemstra, and I. Ruthven, Eds. ACM.

JOCHMANN-MANNAK, H., LENTZ, L., HUIBERS, T., AND SANDERS, T. 2012. Three types of children's infor-
    mational web sites: An inventory of design conventions. *Technical Communication 59,* 4, 302–323.

JONES, R. AND KLINKNER, K. L. 2008. Beyond the session timeout: automatic hierarchical segmentation
    of search topics in query logs. In *CIKM '08*. ACM, New York, NY, USA, 699–708.

JONES, R., KUMAR, R., PANG, B., AND TOMKINS, A. 2007. "I know what you did last summer": query logs
    and user privacy. In *CIKM '07*. ACM, New York, NY, USA, 909–914.

KAIL, R. 2009. *Children and Their Development*. Pearson Education. Pearson, USA.

KAMMERER, Y. AND BOHNACKER, M. 2012. Children's web search with google: the effectiveness of nat-
    ural language queries. In *Proceedings of the 11th International Conference on Interaction Design and
    Children*. IDC '12. ACM, New York, NY, USA, 184–187.

KUHLTHAU, C. C. 1991. Inside the search process: Information seeking from the user's perspective. *JA-
    SIS 42,* 5, 361–371.

KUHLTHAU, C. C. 1999. Accommodating the user's information search process: challenges for informa-
    tion retrieval system designers. *Bulletin of the American Society for Information Science and Tech-
    nology 25,* 3, 12–16.

KUMAR, R. AND TOMKINS, A. 2010. A characterization of online browsing behavior. In *WWW '10*. ACM,
    New York, NY, USA, 561–570.

LIVINGSTONE, S., HADDON, L., GRZIG, A., AND LAFSSON, K. 2011. EU Kids Online Final Report II. The
    London School of Economics and Political Science. http://www.children-go-online.net.

MADDEN, M., LENHART, A., DUGGAN, M., CORTESI, S., AND GASSER, U. 2013. Teens and Technology
    2013. PEW Research Center, The Berkman Center for Internet and Society at Harvard University.
    http://www.pewinternet.org.

MOORE, R. K. AND BOSCH, L. T. 2009. Modelling vocabulary growth from birth to young adulthood. In
    INTERSPEECH. Curran Associates, Inc., USA, 1727–1730.

NAHL, D. AND HARADA, V. H. 1996. Composing boolean search statements: Self-confidence, concept analy-
    sis, search logic, and errors. *School Library Media Quarterly 24*, 199–207.

NEUMAN, D. 1995. High school students' use of databases: Results of a national delphi study. *Journal of the
    American Society for Information Science 46,* 4, 284–298.

OFCOM. 2010. Uk children's media literacy: Research document.

PHAN, N., BAILEY, P., AND WILKINSON, R. 2007. Understanding the relationship of information need speci-
    ficity to search query length. In *SIGIR '07*. ACM, New York, NY, USA, 709–710.

RICHTEL, M. 2011. Children fail to recognize online ads, study says. http://bits.blogs.nytimes.com/2011/04/
    21/children-fail-to-recognize-online-ads-study-says/?hpw.

RICHTEL, M. AND HELFT, M. 2011. Facebook users who are under age raise concerns. http://www.nytimes.
    com/2011/03/12/technology/internet/12underage.html?pagewanted=1&_r=1.

RIEH, S. Y. AND XIE, H. I. 2006. Analysis of multiple query reformulations on the web: The interactive
    information retrieval context. *Information Processing & Management 42,* 3, 751–768.

ROUL, R. K. AND SAHAY, S. 2012. An effective information retrieval for ambiguous query. *arXiv preprint
    arXiv:1204.1406*.

SCHACTER, J., CHUNG, G. K., AND 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.

SMITH, P., COWIE, H., AND BLADES, M. 2011. Understanding children's developments.

SOLOMON, P. 1993. Children's information retrieval behavior: A case analysis of an opac. *Journal of the
    American Society for Information Science 44,* 5, 245–264.

THELWALL, M., BUCKLEY, K., PALTOGLOU, G., CAI, D., AND KAPPAS, A. 2010. Sentiment in short strength
    detection informal text. *Journal of the American Society for Information Science and Technology 61*,
    2544–2558.

TONZAR, C., LOTTO, L., AND JOB, R. 2009. L2 vocabulary acquisition in children: Effects of learning method
    and cognate status. *Language Learning 59,* 3, 623–646.

TORRES, S. D., HIEMSTRA, D., AND SERDYUKOV, P. 2010. Query log analysis in the context of information
    retrieval for children. In *SIGIR' 10*. ACM, New York, NY, USA, 847–848.

TYLER, S. K. AND TEEVAN, J. 2010. Large scale query log analysis of re-finding. In *WSDM '10*. ACM, New
    York, NY, USA, 191–200.

VAN DIJK, E. M. A. G., VAN DER SLUIS, F., AND NIJHOLT, A. 2012. In *Proceedings 4th International ICST Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN 2011), Genova, Italy*, A. Camurri and C. Costa, Eds. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering Series, vol. 78. Springer Verlag, Heidelberg, 139–148.

VETTORI, C. AND MICH, O. 2011. Supporting deaf children's reading skills: the many challenges of text simplification. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*. ASSETS '11. ACM, New York, NY, USA, 283–284.

WEBER, I. AND CASTILLO, C. 2010. The demographics of web search. In *SIGIR '10*. ACM, New York, NY, USA, 523–530.

WEBER, I. AND JAIMES, A. 2011. Who uses web search for what: and how. In *WSDM '11*. ACM, New York, NY, USA, 15–24.

WEBER, I., UKKONEN, A., AND GIONIS, A. 2012. Answers, not links: extracting tips from yahoo! answers to address how-to web queries. In *WSDM '12*. ACM, New York, NY, USA, 613–622.