Usability and Perception of Young Users and Adults on Targeted Web Search Engines

Tatiana Gossen  
Otto von Guericke University  
Magdeburg, Germany  
tatiana.gossen@ovgu.de

Juliane Höbel  
Otto von Guericke University  
Magdeburg, Germany  
juiliane.hoebel@st.ovgu.de

Andreas Nürnberger  
Otto von Guericke University  
Magdeburg, Germany  
andreas.nuernberger@ovgu.de

ABSTRACT

The usability of web search engines is an important factor that influences user experience and correlates with users’ success in finding the relevant information. Currently, there are different search engines online available whose main target audience are children. In this paper, we investigate the differences between children and adults in terms of usability and perception of targeted search engines, i.e. search engines designed specifically for that audience. To this end, an eye-tracking study was conducted to compare children’s and adults ‘ search behavior and perception of search interface elements on search engine results pages (SERPs) during an informational and a navigational search with a standard search engine and a search engine for children. We identified differences in the information-seeking behavior and perception of search engines SERPs between children and adults. Based on these findings we propose criteria on how to design search user interfaces that are more appropriate for children.

Categories and Subject Descriptors

H.5.2. [User Interfaces]: Evaluation/methodology; H.3.3. [Information Search and Retrieval]: Search process

Keywords

Search engine, user study, eye-tracker, children

1. INTRODUCTION

These days, Internet usage has no real age limit. Many children have access to the Internet and explore the Web from a young age on. The German study [26] reports that about 60% of the children of ages six to thirteen use the Internet, and 70% of those use search engines. Children use the Internet for entertainment, e.g. online games. Furthermore, they also use the Internet to research for their school activities [13]. More than half of the children search the Internet predominantly alone [26].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

BIX’14, August 26 - 29 2014, Regensburg, Germany
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-2976-7/14/08...$15.00.
http://dx.doi.org/10.1145/2637002.2637007

Therefore, there are special search engines available that aim at supporting children during information acquisition. The majority of the engines have a colorful designed start page to attract children’s attention (e.g. kidrex.org, blinde-kuh.de, helles-koepfchen.de, quinturakids.com). In every other aspect, existing web search engines for children have the similar design to common search engines [11]. As these search engine do not address children’s cognitive, fine motor and other abilities, they have some issues in usability [11]. To increase the usability of search engines for children, understanding of users’ search behavior and interface perception is essential and can be achieved with the help of user studies. Previous research shows differences between children and adults in their search behavior. However, these findings were mainly based on observational studies, e.g. [2], or log file analysis, e.g. [13], and do not provide details about user perception of the search engines’ interface elements during information seeking. To complement existing findings eye-tracking devices can be used.

Eye-tracking provides information about users’ line of sight at any given time. In the context of a web search, this means that information can be gathered on what web interface elements caught users’ attention (fixations), for how long (fixation duration) and in which order (scanpaths). According to Granka et al. [15] “a fixation is generally defined as a spatially stable gaze lasting for approximately 200-300 milliseconds, during which visual attention is directed to a specific area of the visual display”. The so-called Strong eye-mind Hypothesis [20] states that a displayed item that is fixated is also being thought about. However, this assumption has some limitations, i.e. the hypothesis holds only if the user’s current task requires information from the visual display to be encoded and processed [19]. Thus, we assume that this hypothesis holds in case of an information seeking task. Nevertheless, duration of the gaze provides only an upper bound for the duration of cognitive processes [19]. To summarize, eye-tracker data is very important in order to increase the usability of search user interfaces (SUIs) and to design novel search engines for children.

In this paper we describe the design and results of an eye-tracking user study with children of primary school age and adults to study their seeking behavior and search engine perception during informational and navigational search (in terms of Broder’s taxonomy [4]). The purpose of informational search is to find information about a topic assumed to be available on the web. Children tend to employ informational search [13]. Adults most frequently employ navigational search with the immediate intent to reach a particular
website that the user has in mind [13]. Therefore, we study both search variants. We also study the information seeking behavior not only on a standard search engine (Google), but also on a web search engine for children (Blinde-Kuh.de).

In our work we study the search behavior of children in third and fourth school grade. At this age, children already have sufficient reading and writing skills to perform the web search, but they are only superficially familiar with web search engines [26]. Therefore, they are an appropriate user group to study how intuitive search engines are. Druin et al. [10] also call these children “developing searchers” who have challenges with spelling, typing, query formulation and results interpretation.

In summary, the contributions we present in this paper are threefold: 1) Using an eye-tracking user study, we analyze the differences in information seeking behavior between children and adults during informational and navigational search. 2) Furthermore, for these two user groups, we compare the information seeking behavior on two different search engines, i.e. Google and a web search engine for children. 3) Our results provide a deeper understanding of children’s search behavior and perception of search interface elements, so that better search user interfaces for children can be designed in the future. To our knowledge, this is the first comparative study between children and adults, during both informational and navigational search, on both standard and children’s search engine.

2. RELATED WORK

In this section, we summarize existing eye-tracking user studies about user information seeking behavior on web search engines. A more general introduction to the usage of eye-tracking in online search is given in [15].

Granka et al. [16] studied the search behavior of 36 adults when using Google during a task-oriented search given both navigational and informational tasks. The results show that the users spent almost the same time to review the first and the second result. After the second result, the fixation time decreases for each following search result. Results ranked 6 to 10 receive approximately equal attention, which can be explained by the fact that these results are placed under the scrolling line. Therefore, Granka et al. consider the rank of a search result to be less relevant when the user starts to scroll. Furthermore, they note that users seldomly click on the seventh result which lies directly below the page break and the scroll border. Users who select the search results with a low rank also tend to read more textual snippets of the search results. The scanning behavior of users is rather linear as users tend to explore the search results list from top to bottom.

Rele and Duchowski [31] studied the search behavior of 15 adults during a task-oriented search given both navigational and informational tasks. In particular, they compared a list interface, where the results were listed vertically, with a tabular interface, where each element in the list interface had a corresponding column. The study results show no significant difference in performance or average fixation duration between the two interfaces. The number of fixations in the text descriptions of the search results differs significantly for different types of search tasks: There are more fixations present in the textual snippets given a navigational search task than there are in informational ones. Rele and Duchowski explain this fact as an extensive reading of the snippet content because the navigation tasks are more difficult to solve. An interesting result of the study is that users rather tend to scan within columns and are less likely to shift gaze between the columns. The probability of gaze shift between the same category of result snippets, e.g. from the URL of the first result to the URL of the third one, is larger on the tabular interface.

These two studies provide a detailed quantitative analysis of the distribution of visual attention on search engine results pages (SERPs). There also exist studies about visual attention on other aspects of web searches, such as reading. According to Nielsens and Pernice [27], users’ reading behavior across many different web sites and search tasks is almost consistent. It follows a reading pattern in the shape of an “F”: in the upper section of the website, users tend to read in the horizontal direction first. This generates the upper bar of “F” shape. After that, the eyes move a little in the direction of the web site bottom followed by eye movement in a second horizontal direction that, however, covers a smaller area than in the first one. Thereby, the lower horizontal bar of the “F” shape is formed. Finally, the user scans the left web page side in the vertical direction. This scanning can be done slowly and methodically, which appears as a continuous strip on the heat map. On the other hand, the eye movements can also be fast, which leads to a spotty heat map. Nielsen and Pernice [27] also found some variance in visual strategies, e.g. the lower horizontal bar of “F” can be longer than the upper one. This deviation occurs in case the title of the second search result is longer than the title of the first one [8, 27].

Previous user studies with adults also explored different aspects of web search behavior. For example, Cutrell and Guan [7] studied the influence of the snippet length on users’ search behavior. Adding extra information in the snippets improves the performance significantly in informational search tasks, while the performance in navigational search tasks is best if the snippets are short. Lorigo et al. [25] analyzed the influence of gender and task type on Google web searches. The users in the study spent more time solving an informational task than a navigational one, the men explore more search results than the women. Lorigo et al. [24] surveyed the use of eye tracking to study users’ online search, in particular how users view the ranked results on SERPs and the click-through behavior. They also conducted a study to compare users’ viewing pattern on Google and Yahoo! search engines. No differences in terms of search performance and eye movements behavior were observed between the two search engines. Pan et al. [28] studied the influence of search results’ relevance and its ranking position. Users expect the first result to be the most relevant and click on it even in case the ranking order in the result list is reversed. Aula et al. [1] studied user’s style to evaluate search results. Two types of evaluators were found, economic and exhaustive evaluators. Evaluators with economic style decide faster about the next move and based on less information than exhaustive ones, therefore being more efficient in search. Buscher et al. [6] studied contemporary search engine elements such as ads and “related” searches. They found that users’ attention to ads depends on both task type and ads quality. Joachims et al. [17] evaluated the reliability of implicit feedback generated from click-through data and suggested that users clicks are influenced by the relevance of the results, but are biased by the users’ trust in
the quality of the retrieval function and the result set.

There are diverse user studies about children’s information seeking behavior, e.g. [2](seventh grade), [3](ages 9-12), [18](ages 8-12), [21](ages 8-10), [23](sixth grade). An overview of these studies is given in [14]. Based on these studies, children require emotional, language, cognitive, memory, interaction and relevance support. In contrast to adults, children have a limited domain knowledge, which causes difficulties to formulate a query in case of a keyword-oriented search. Furthermore, using a keyboard while typing, children often do not spot spelling mistakes. Compared to adults, children have a different navigational style, called a loopy browsing. Children click, repeat searches and revisit the same result web page more often than adults. This behavior can be a sign of children’s cognitive overload during the search. Children can get easily frustrated if they do not find relevant results, do not understand the search engine output or if a failure emerges. The fact that children also have difficulties to evaluate the relevance of retrieved documents to their information needs aggravates this.

To our knowledge there exist only one eye-tracking study about children’s perception of SERPs. Dinet et al. [8] conducted a user study with young users in the age from 10 to 17 and observed user behavior on Google SERPs. They suggest that keyword-highlighting and user domain knowledge both influence the users’ visual strategies. The results of this study show that the visual exploration of young children is influenced by highlightings, regardless of prior knowledge of the child and the location of the highlighted terms. The search strategies of older children are affected by highlighting only when the subject of the Web search is unknown. Dinet et al. describe four different search strategies of children: 1) F-shaped strategy, 2) exhaustive strategy, where users read all the information in the result document representative (surrogate) before clicking a search result, 3) cued visual jumps, where users after reading one highlighted keyword jump to the next one, 4) F-inverse strategy, where users examine the bottom of the SERP first. These four strategies are present in all age groups of the users of this study. However, younger children tend to use the “cued visual jumps”-strategy, whereas older children often extensively read search result’s title, snippet and URL. Older kids sometimes also use the “cued visual jumps”-strategy, especially in case they are not familiar with the search topic or have little knowledge about it.

3. USER STUDY

In this section we describe the design and results of our user study. In our study we used Google (Fig. 1(a)) and the German search engine for children Blinde-Kuh.de (Fig. 1(b)). Blinde-Kuh.de is often used in German primary schools. Furthermore, the search user interface of Blinde-Kuh.de differs from Google’s. This makes it an appropriate search engine for children to study. Blinde-Kuh.de (BK) offers a keyword search (A) along with navigation in different categories provided on the left side (B). These categories lead to informational pages about the corresponding topic. The search engine outputs at most ten results per page as a vertical list. Search results are separated through boxes in contrast to Google where separation is done by whitespace. Each BK surrogate contains a picture (C), textual summary (D), rank (E), information about the result’s category (F), the target age group in categories “S”, “M”, “L”, “XL” for children from six to thirteen (G) and does not use any keyword highlighting. To see the BK search result’s page, one can click on the title (H), picture or a link below (I).

The main research questions of this work are: 1) Are children more successful with BK than with Google? 2) Do children prefer BK over Google? 3) Do children and adults have different perception of web search interfaces? 4) Do children’s and adults’ employ different search strategies?

The Within Participants Design was used in our study. In
particular, we used a two-stage experimental design with two factors: the type of web search engine (standard web search engine and children’s web search engine) and the search task type (informational and navigational). We also applied a Latin square design, where each participant interacted with both search engines, but in a different order to avoid biases due to usage order, tiredness, etc. We used two informational (Info) and two navigational (Nav) search tasks. Each participant was given one navigational and one informational task. However, due to differences in the provided results between the two search engines and to construct appropriate search scenarios, different search tasks with similar complexity were used for different search engines. The organization into groups is shown in Table 1.

Procedure: We tested one participant at a time. Using a structured pre-interview, we gathered the user’s demographic data and Internet experience. Users were told that first they would receive search results for a query provided by the study supervisor. Thus, participants began on the same SERP, but were allowed to proceed in any way they chose. After that, the eye-tracker Tobii Eyetracker T60 was calibrated that is integrated in a 17” monitor. Then, each subject received a search task to solve within 10 minutes for each search engine. After each search task, we also asked the participants about their own assessment of search task difficulty and how they like the search engine’s results pages (SERPs). At the end, a structured post-interview about user preferences in search engines was performed.

The search tasks and corresponding initial queries used in the study are presented in Table 2. The scope of a task is designed so that it can be solved in a reasonable amount of time, but the solution is not trivial. Initial queries were selected such that they had at least two SERPs, and that the correct answer could be found in one of the first 20 results. For navigational tasks there was only one correct result web page. The initial SERPs contained significant SUI elements. We ensured the search consistency in order for the search conditions to be the same among all participants. In specific, the same SERPs were presented to the users using initial queries. As SUIs could be altered by the search engine owners, we fixed the first two SERPs for the users using initial queries. In specific, the same SERPs were presented to the participants. In specific, the same SERPs were presented to the participants. As SUIs could be altered by the search engine owners, we fixed the first two SERPs for the users using initial queries. The resulting HTML documents had the same characteristics as dynamically generated Google SERPs. After each session, the browser history was automatically deleted to avoid highlighting of previously clicked search results. All browser cookies were automatically deleted to disable the personalization of search results.

Participants: The user study was conducted in February 2013. We collaborated with a primary school in Biederitz, Germany. 14 children participated in the study. They were between eight and eleven (9.29 on average, \( \sigma = 0.73 \)), 64% were boys. 43% from third and 57% from fourth grade. 43% of the children use the Internet once per week. 64% of the pupils use the Internet without any supervision, whereas the rest of the participants are accompanied by their parents or older siblings. 57% of the young participants use the Internet mostly to play online games and to watch videos. 36% also use the Internet to search for information. 7% use the Internet only to search for information, for example, to do their homework. More than 80% are familiar with Google and 46% also know BK. Children like both Google and BK because they can provide “a lot of information”.

Adults were recruited from an academic and school context. 17 adults also participated in the study. They were between 22 and 59 (29.81 on average, \( \sigma = 8.93 \)). 65% were male. 70% of adults are students in computer science or working in the IT sector. This introduces one side effect. These adults can be considered as search experts, while the children are novices in web search. All the adults use the Internet every day without any supervision. 18% use the Internet only to search for information. The rest also use the Internet for other activities such as chatting etc. All the adult participants usually use Google to search for information. 29% of the adults use other search engines such as Yahoo, Bing or DuckDuckGo along with Google. Adults told us that Google is “concise and user-friendly”, it offers “a good balance between speed and quality of search results”.

4. STUDY RESULTS

Search effectiveness: The effectiveness of a user’s search describes the degree of success in finding the relevant information or the requested website. We consider not only the fact of finding the result, but also whether a participant required help from the study supervisor telling him what is possible to do next. Table 3 shows the calculation of success scores based on two variables, finding the right solution and required help. The best result (4) is achieved when the task is solved without any hints.

Fig. 2 provides details about participants’ success scores. The participants are on average equally successful on both Google and BK, during both informational and navigational search. The data for the children are more scattered than for the adults. 50% of the children are “successful” in search or

<table>
<thead>
<tr>
<th>ID</th>
<th>Search task (Initial Query)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info-1</td>
<td>How many animals does the zoo in Magdeburg have? (Zoo Magdeburg)</td>
</tr>
<tr>
<td>Info-2</td>
<td>What is the name of the largest Saturn moon? (Saturn Monde)</td>
</tr>
<tr>
<td>Nav-1</td>
<td>Find the homepage of the photographer Michael Jordan. (Michael Jordan)</td>
</tr>
<tr>
<td>Nav-2</td>
<td>Find the online-game page of ZDF Tivi portal. (ZDF Tivi)</td>
</tr>
</tbody>
</table>

Table 2: Search tasks and corresponding initial queries used in the study.
have a lower score. This shows that children on average can solve a search task, but not without hints to continue the search. The adults are “very successful” in search and do not require any help. There was only one failure while using BK: This adult did not solve a search task within 10 min, probably because it is not a standard search engine for adults. The child who achieved score 1 is a third grade pupil and uses the Internet only once a month and mainly to chat with friends. The child who achieved score 2 is a fourth grade pupil and uses the Internet once a week but under parents supervision and also mainly to chat with friends. Their failures can be due to little search experience. The Mann-Whitney-U-test showed a significant difference between children’s and adults’ success scores \(p < 0.001\). The Wilcoxon signed-rank test showed no significant difference between children’s search with Google and BK, between children’s search during navigational and informational task. No significant differences were found for adults.

**Search efficiency:** We measured the time required by participants to solve the search tasks. As mentioned above, the time for each task was limited to 10 minutes. The results are given in Fig. 3. The children needed on average four times longer to solve a search task (298 seconds versus 76 on average). The t-test shows a significant difference between children’s and adults’ search times \(p < 0.001\). This can be explained by the fact that children read slower than adults do [5]. Children also have more difficulties to determine the relevance of results [14, 18], which also results in longer times for children.

The data for the children are more scattered than for the adults. On average, the children required more time to solve informational search tasks. This is consistent with the fact that more reading within web documents is required to solve an informational task. Navigational tasks are faster to solve because the hints in the title or URL are usually sufficient to successfully navigate to the required web page. The data for children are more scattered for the search with BK than with Google and for informational search than for navigational one. There is also a significant difference between adult’s search times with BK and with Google \(p < 0.05; 114\) seconds with BK and 45 with Google on average), however no significant differences were found for children. The difference for adults can be due to the lack of familiarity with the children’s search engine. In addition, the very colorful and visually overloaded design of BK frustrated the adult participants who preferred the simple Google user interface.

**Difficulty estimation:** After each search task the participants were asked to provide information about how difficult the task was using a five-point Likert scale. We used a five-point smiley scale for children called smileyometer [30]. The children estimated their searches as easy on average, while adults found the searches even easier, especially when using Google. There is no obvious difference in children’s difficulty estimation between BK and Google.

**Search engine preference:** According to the post-interview, all the adults and 58% of the children prefer to use Google. The children like the SERPs of both search engine equally well. The adults rate the SERPs of BK as moderate and Google’s as good on average. The children liked that BK is “lovely colorful”, has “not only text, but also pictures” and offers “a lot of information for children”. One child positively mentioned that one “can not only search but also use many other features” of BK (e.g. many categories). Two children found the query suggestions of Google to be helpful. One child had difficulties to find the search box on Google. Children also found it to be “great that the (Google) SERP has videos”. The adults told us that Google’s SERPs are “well-structured” and “clearly arranged”. They liked the query completion function of Google. However, especially for navigational search, the textual summaries were too short. Adults told us that the SERPs of BK are “very unclear and contain many elements that distract from the actual search”, “chaotic” and have an “overloaded layout”. Furthermore, according to the adults opinion BK SERP “images contain unclear information”. The adults positively mentioned that the SERPs of BK offer meta information

<table>
<thead>
<tr>
<th>Calculation of success scores:</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finding the right solution</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Required help</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 3: Calculation of success scores: 4 “very successful”, 3 “successful”, 2 “less successful” and 1 “not successful”.

Figure 2: Search effectiveness: boxplots show participants success during the search. Children (C.) are on average “successful” in search and adults (A.) are “very successful” on the scale from 4 - “very successful” to 1 - “not successful”.

Figure 3: Search efficiency: boxplots show the search time to solve a given search task. Children (C.) require on average four times longer than adults (A.) to solve a search task.
about search results such as preview, categories etc.

Heat maps: In order to analyze the participants’ perception of SERPs we created heat maps. Unfortunately, not all the eye-tracking data is usable as children are very agile and some of them moved too close to the monitor during the session for the eye-tracker to capture their eye movements. Our analysis is based on 11 young users and 17 adults. Fig. 4, 5, 6 and 7 show several aggregated heat maps both for children and adults, i.e. accumulated number of fixations from all the selected test persons. Regions in red indicate a high number of fixations, whereas green regions indicate very few fixations, with varying levels in between. The heat maps show a difference between the children’s and the adults’ search behavior.

On both search engines, the heat maps for children contain relatively closely spaced circular color-highlighted areas that are spread across all titles of the search results list. This scan pattern indicates a partial reading of the search results’ titles. Furthermore, there are visual jumps from a highlighted word in the title to the next one available. Due to this reading style, relevant information can be overlooked, so that children often navigate to the second SERP. Our findings are consistent with Dinet et al. [8] who found a tendency towards a cued visual jump strategy of younger children. However, we also found that children tend to scan the whole result list using jumps between highlighted words, especially in titles.

During navigational search the children paid more attention to snippets than during informational search for both Google and BK. This can be explained by the fact that navigation tasks require an extensive reading of the snippet content [31]. The adults only fixated the first search results. The fixation areas get smaller for results ranked lower in the list. This pattern is consistent with the F-shaped strategy [27]. In case the answer is not found within the first search results, the adults reformulated the query. This behavior is similar to the depth-first strategy of processing search result lists described in [22].

Children’s strategy of processing search result lists is more like the breadth-first strategy [22]. Young users exhaustively explored all the results in the first SERP and used the navigation buttons between results pages to continue further examination. We assume that navigation is easier for children than query modification since children’s low capacity for abstraction [29] makes it difficult for children to create their own queries. Children reformulated the query mostly on the second SERP. Overall, there is no significant difference in the number of query reformulations (children: 1.69 on average, \(\sigma = 1.702\); adults: 1.18 on average, \(\sigma = 1.629\)).

On both search engines, children paid more attention to media elements such as pictures and video elements embedded in a search result than adults. On Google, children’s heat maps contain a large cluster of fixations in the middle of SERPs, whereas adults’ heat maps show a high number of fixations in the top of SERPs. The Google search engine provides query suggestions placed on the bottom of the page. However, only children paid attention to query suggestions and adults did not.

In contrast to Google, the BK search engine also provides thumbnails for nearly each search result and uses no highlighting of keywords. The heat map of BK shows a Γ-shaped scanning strategy of children within a search surrogate as children tend to scan the title and look at the web page picture. Adults did not fixate thumbnails as much as children do. Spotty and closely spaced colored areas in the titles of search results indicate a partial reading of children, the same as for Google. Children fixated informative words, e.g. “television broadcasting” or “tivi online” in the title, but also in the snippet and the link of a search result. The links in the search results, for example the link with the words “Next to Space Agents”, were barely perceived by children.

In opposite to adults children looked at the navigational menu with categories on the surface of the BK. Some children even clicked on the category “Games” that is related to their navigational task to find the online-game page of ZDF Tivi portal. Both children and adults had difficulties to determine what parts of BK result surrogates are clickable. Probably, the box around a surrogate let them to believe that the whole area is clickable. Fig. 8 shows a click on a textual summary that is not clickable.

Relative fixation duration: Fig. 9 shows the distribution of the fixation duration depending on the rank of a search result\(^6\). The adults looked at the first result for the longest.

\(^6\)The SERP for Google’s informational search had only seven results. The first result contained shortcuts to the underly-
There is a trend that the higher the results rank the less time the adults spent to review it. Overall, children’s fixation time did not correlate with the search results’ rank, but rather with search results’ relevance clues such as highlighted words in the snippet (see the 4th result for informational search in Google), informative words in the title (see the 9th and 10th result for information search in Blinde-Kuh) or embedded media such as Youtube videos (see the 7th result for informational search in Google).

Fig. 10 shows average fixation duration for different surrogate elements of Blinde-Kuh such as link, title, snippet and picture (BK surrogates do not contain URL). The children spent more time fixating pictures, especially in navigational search, than the adults. The children fixated the title longer during an informational search and the picture during a navigational one. The adults fixated the title, the snippet and the picture equally long during informational search, while the title was fixated longer during navigational search.

Gaze plots: A gaze plot is a map which shows eye movement sequence, i.e. gaze fixations on a webpage and in which order they occur. According to Lorigo et al. [25] there are three types of scanning sequences or scanpaths: nonlinear scanning, linear scanning, and strictly linear scanning (see also [15]). Nonlinear scanning implies that the search results are viewed not in a rank induced sequence order, but randomly. Linear scanning means that the search results are viewed sequentially, i.e. a result of rank n is not viewed until all results of a smaller rank have also been seen. However regressions to previously visited results are possible. Strictly linear scanning obeys the rank order and has no regressions to earlier surrogates.

We manually reviewed the individual gaze plots obtained in the study. Scanpath sequences of the adults are relatively short (they only review the first results), while children’s scanpath sequences are long and appear chaotic (see Fig. 11). During the search with Google, all young users in our study employ the nonlinear scanning strategy, reviewing the results in a random order together with returning to results which they already viewed. The adults’ reading is more linear (40% of adult searches with Google are strictly linear, 60% of adult searches with Google are nonlinear). The results for adults are consistent with [25], who found that adults skip some results while reading. During the search with BK, both the adults and the young users in our study tend to employ the linear scanning strategy: reviewing the results in a sequentially order, but also returning to results which they already viewed. 80% of children’s searches with
Figure 9: Average fixation duration for the ten search results.

Figure 10: Proportion of average fixation duration for different surrogate elements of Blinde-Kuh such as link, title, snippet and picture. Surrogate that contain no picture are marked with *.
BK are linear and 20% nonlinear. 67% of adults’ searches with BK are linear and 33% nonlinear. This difference to Google can be explained by the fact that the BK results list is longer in term of required screen place. Therefore, the participants have to scroll more to see further results which results in a more sequentially order.

5. DISCUSSION

Based on the findings of this study we can answer our research questions as follows: Both search effectiveness and efficiency indicate that for children there is no significant difference between Google and Blinde-Kuh, both in informational and navigational search. Our results are consistent with Jochmann-Mannak et al. [18] who found that the children did not perform better on interfaces designed for them than on Google. Children tend to prefer BK a little less than Google. The children’s preference of Google search engine is consistent with Druin et al. [9] who found that children’s perception of the Google was quite positive. The large search times of children in comparison to adults indicate the need to mitigate children’s difficulties in reading that can be done by texts simplification. We also found that the very colorful and visually overloaded design of BK frustrated the adult participants and the frustration even lead to a decrease in performance. Therefore, adults should not use search engines targeted for children. We found differences in the perception of search engines SERPs between children and adults. User interface elements such as thumbnails and embedded media attracted children much more than adults. Though BK provides pictures, they are apparently not well selected as they do not lead to an improvement in children’s performance. Therefore, we suggest that those elements be only used in SERPs if they provide better relevance cues so as to improve children’s reading flow.

Furthermore, there is a difference in distribution of fixation duration between adults and children. There is a negative correlation between the rank of a search result and the fixation duration of adults. However, children’s fixation time rather correlates with search results’ relevance clues or embedded media. This also confirms the need to provide better relevance cues in media and texts for children. The children only partially read the results’ snippets. Therefore, it may be unnecessary to present a long text in each snippet of a search result. Instead, snippets should be short and provide relevant typographical cues to avoid frustration when reading long texts and to help children in finding relevant information. The children also used navigational elements and at least looked at the menu with categories, while adults did not. Therefore, these elements should be part of a search engine and support children who have difficulties to formulate a new query.

Adults and children employed different search strategies. Children’s strategy for processing search result lists is similar to breadth-first strategy: They exhaustively explored all the results in the first SERP. Adults only viewed the first results and then reformulated the query. There is also a difference between adults and children in result scanning style on Google: children viewed the results in a random order, while adults tend to scan the results in a rank order.

6. CONCLUSIONS

We conducted an eye-tracking study to compare children’s and adults’ information-seeking behavior on SERPs of a standard search engine (Google) and a search engine for children (Blinde-Kuh), both during informational and navigational search. We found that children are equally successful using both search engines and tend to prefer Google over Blinde-Kuh. Adults are also more successful using a standard search engine and do not like the overloaded interface of the children’s search engine. Children and adults have different perceptions of the SERP elements. The identified differences can help to derive criteria for designing better search user interfaces for children. The study shows strong evidence of children’s difficulties to determine the relevancy of a search result. Children use a breadth-first-like search strategy examining the whole result list, while adults only examine the first top results and reformulate the query. To support children in relevancy estimation, new solutions should be provided in the future, for example, by designing other representatives of a website that matches children’s view. The study had some limitations in the sampling nature and sampling size. The young participants of the study were from a more generic population (a school class) than the adults (related to IT). Despite the fact that our results for adults are consistent with other research conducted with larger and more generic user groups, e.g., the F-shaped strategy of adults [27], complimentary user studies with both
larger sample and more general adult population could lead to additional findings.

7. ACKNOWLEDGMENTS

We would like to thank all participants of the user study. This work was partially supported by the Transregional Collaborative Research Centre SFB/TRR 62 “Companion-Technology for Cognitive Technical Systems” funded by the German Research Foundation (DFG).

8. REFERENCES


