Right Way, Right Time: Towards a Better Comprehension of Young Students' Needs when Looking for Relevant Search Results

Monica Landoni
monica.landoni@usi.ch
Università della Svizzera Italiana,
Lugano
Switzerland

Mohammad Aliannejadi
m.aliannejadi@uva.nl
University of Amsterdam, Amsterdam
The Netherlands

Theo Huibers*
t.w.c.huibers@utwente.nl
University of Twente, Enschede
The Netherlands

Emiliana Murgia
emiliana.murgia@unimib.it
Università degli Studi di
Milano-Bicocca, Milano
Italy

Maria Soledad Pera
solepera@boisestate.edu
PIReT - Dept. of Computer Science,
Boise State University, Boise
United States

ABSTRACT

We report on the exploration we conducted to understand better children's needs for the design of Search Engine Result Pages (SERP) that can help them notice relevant resources when performing online inquiry tasks in a classroom context. We analyse children's interactions with traditional and emoji-enriched SERP and look for trends linking children's engagement with SERP and search success (based on experts' assessments). We also identify areas remaining to be interpreted and considered in future studies. On this mixed ground, we discuss the complexity of this design space and the need to bypass the one-size-fits-all approach in favor of adaptive SERP to cater to children's different and ever-evolving skills in searching and recognising useful results if we aim to support learning.

CCS CONCEPTS

• **Social and professional topics** → **Children**; • **Information systems** → *Search interfaces*.

KEYWORDS

search, children, classroom, user modeling, support, scaffolding

*Co-founder of Wizenoze.

1 INTRODUCTION

As members of the "Google generation," [17], young users born into the digital age have only ever needed to query a search engine to get "some kind of answer, even if it may not be the best one" [10]. Being familiar with search tools, however, does not equate to knowing how to use them well. Ever since the Puppy-IR project, which took place between 2009 and 2012, it has been a cradle of many follow-up research projects [6, 11], researchers and practitioners have continued to move forward towards designing algorithms and interfaces that can support children's information-seeking practices [2, 5, 7–9, 12]. These solutions, however, often respond to the stereotypical behaviour and cognitive abilities of children at different ages, and not to children's specific needs and task objectives, which is a must when using computers in the classroom to support learning [18] and the focus of the quest described in this paper.

Aliannejadi et al. [1] shared preliminary evidence showcasing that Search Engine Result Pages (SERP) enriched with emojis can better support the completion of information discovery tasks common to the classroom context. However, outcomes from their work also reveal that (i) which emoji-enriched SERP should be used in the classroom and (ii) when children actually need the extra support that enriched interfaces can provide are two concerns that remain uncertain. We seek to advance this area of study by digging deeper into the data collected by Aliannejadi et al. [1] via a user study with children (ages 10 - 11) in two primary five classrooms to observe their search behaviour. Besides search logs, data includes responses to a post-task questionnaire eliciting children's preferences for using traditional or emoji-enriched SERP to complete search tasks as well as their competence with search and ability to locate information online. We augmented the data with expert assessment on search success (i.e., grades on submitted assignments). We leverage this enriched data in generating a comprehensive picture of the factors contributing to the overall search experience. In turn, it will inform the design of adaptive SERP that can offer scaffolding to support children searching in the classroom [4].

Our empirical analysis aims to unveil connections (or lack thereof) across children's preferences for interfaces, their perceived understanding of the search process, behavioural search traits, and search

success. We outline three research questions to drive our scrutiny of whether and how young searchers take advantage of the support enriched-SERP can provide: **(RQ1)** Does interface choice impact search success? **(RQ2)** Does efficacy perception align with search success? **(RQ3)** Do search behavioural traits impact search success?

Findings from our exploration, while preliminary, reveal a clear need to move beyond a one-size-fits-all user model for children and instead consider the many facets that can define the type of help children need from search tools to successfully utilize them to support learning. Outcomes from our work also evidence the complexity of the design space for producing search tools to be used by children in the classroom and the need for further investigations. Informed by a combination of what we discover as emerging trends and what we identify as areas requiring further examination, we set groundwork pertaining to the design of new user studies that specifically account for a more articulated and representative typology of young searchers and their needs.

2 DATA AND EXPLORATION SET UP

For our exploration, we turn to the **data** collected by Aliannejadi et al. [1]: search logs capturing interactions of 31 children (ages 10-11, 17 boys and 14 girls) in primary five Italian-speaking classrooms in Italy and Switzerland, in addition to responses to a post-task questionnaire. As detailed in [1], data was collected due to a study during which children completed an information discovery assignment related to a common topic in primary five classrooms–Ancient Rome–as part of regular class instruction. The assignment included 12 questions (search prompts); children had to use a given search tool to locate online resources that would help them answer these questions. During the course of this assignment, children were exposed to three different interfaces (in random order but ensuring a uniform distribution among alternatives). Interface `No Emoji` simulates a traditional search engine interface, whereas `Tick/Cross` and `Light Bulb` interfaces are enriched with emojis that serve as cues for relevance or non-relevance of results included in a SERP. Upon completing the assignment, children responded to a post-task questionnaire, where they identified their favourite interface, the one they thought was the most helpful, and the one they would recommend to their peers for future use. They also noted their level of expertise using search tools and whether they thought they have found what they were looking for (distribution of responses shown in Figure 1). Teachers acted as facilitators in setting up the search prompts and explaining to children the steps necessary to accomplish them. The study, which ran completely online due to the COVID-19 emergency, was approved by the ethical board of the respective schools.

With the help of an expert educator, we augmented the data [1] by assigning a **grade** to submitted assignments. We treat this grade as a proxy measure of **search success**, i.e., the degree to which children take advantage of interfaces to locate useful resources and then extract the information needed to complete their assignment on Ancient Rome. A teacher scored each of the 12 submitted answers, which were averaged per child to yield the respective overall grade [0-100]. For analysis purposes, we group children based on their search success: High (grade between 90 and 100),

	No	A little	Yes	A lot
Q1: Did you find relevant info?	3	9	7	7
Q2: Search Experience	0	10	11	5
Q3: Interest in Search	0	3	16	7

Figure 1: Responses to the post-task questionnaire; cells capture number of children.

Medium (grade between 75 and 90), and Low (grade below 75). Each group included 10, 7, and 14 children, respectively.

3 ANALYSIS AND DISCUSSION

We analyze children's overall search experience based on implicit behaviour, explicit perception, and expert assessment. Our exploration is primarily inspired by the important educational concept of *scaffolding*: "the kind and quality of cognitive support which an adult can provide for a child's learning, which anticipates the child's own internalisation of mental functions" [16]. Thus, we pay special attention to the different performance trends that emerge and, therefore, the type of support that search interfaces could and should offer.

Does interface choice impact search success? We start our exploration using the FHAB lenses presented in [1] to capture the user experience supported by varied interfaces from different viewpoints. These lenses are: **Favourite** (interface children liked most); **Helpful** (interface children would use in the future for curriculum-related information discovery tasks); **Advised** (interface children would encourage their peers to use); and **Best Performing** (interface that led children to click on known relevant results).

We examine how enriched interfaces facilitate the completion of information discovery tasks both in *theory* (based on children's preferences) and in *practice* (based on children's interactions with interfaces recorded in search logs in addition to expert assessment of assignments). We expected children to prefer enriched interfaces and thus achieve the highest search success due to the cues afforded by the emojis. As shown in Figure 2(a), children favoring `Light Bulb` achieved the highest search success using the `Light Bulb`-enriched interface to look for resources. Surprisingly, this trend is not sustained when favouring `Tick/Cross`; it appears that aesthetics factors are not necessarily correlated with support offered. However, we did not find significant differences in children's grades regarding the interface they favoured. As depicted in Figure 2(c), children who found emoji-enriched interfaces helpful obtained higher grades using them, as opposed to the `No Emoji` counterpart; the same is true for children who claimed the traditional text-based interface helped them the most. The helpful interface had a significant effect on children's grades (two-way ANOVA, $p < 0.05$). Similarly, but not significantly, children fare the best when taking advantage of the interface they would advise their peers to use
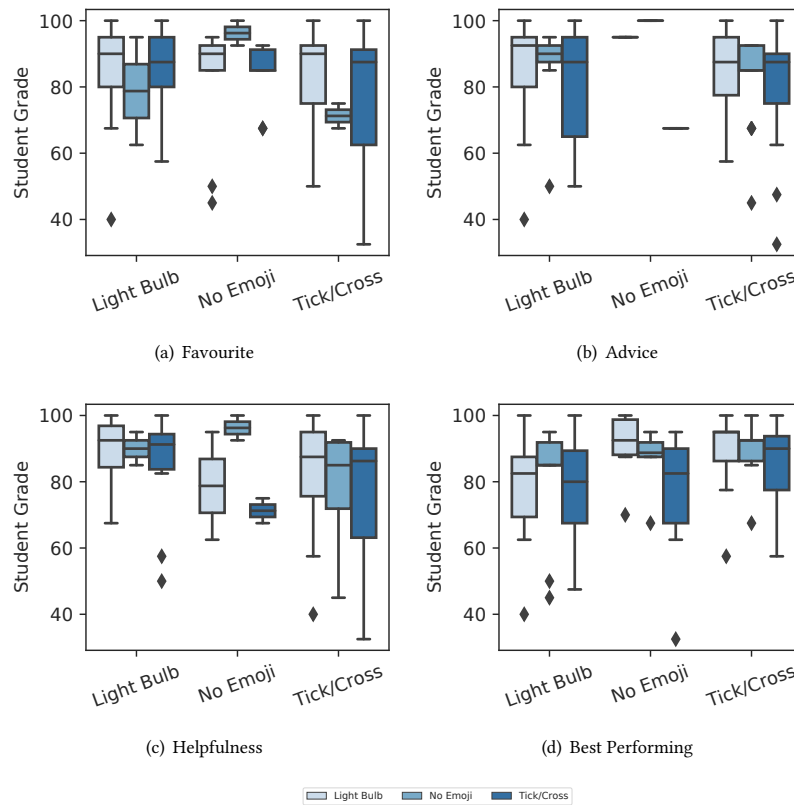
(a) Favourite

(b) Advice

(c) Helpfulness

(d) Best Performing

Light Bulb   No Emoji   Tick/Cross

Figure 2: Exploration based on FHAB lenses and expert assessments. The x-axis indicates the selected interface based on the corresponding lens.

(Figure 2(b)). It can be appreciated from Figure 2(d) that the interface that prompted children to click on more useful resources was not always the one that leads them to the highest grades (two-way ANOVA, $p < 0.05$). For example, children that explored more relevant results while using the Light Bulb-enriched interface obtained higher grades on inquiries they responded to based on their interactions with the traditional interface, i.e., No Emoji. From these results, it becomes apparent that none of the proposed interfaces consistently steer children to successfully completing inquiry tasks. In fact, several students obtained high overall scores even when using a traditional interface. This evidences a knowledge gap regarding user traits that can inform the type of interface support they need, further emphasizing that a 'generic' enriched interface cannot be the solution for all.

Does efficacy perception align with search success? To answer RQ2, we turn to children's perceived efficacy on their abilities to search in general and to successfully complete online inquiry task, in this case on Ancient Rome (via post-task questionnaires); which we compared and contrasted with their search success (via expert assessment).

Children who enjoy searching ('yes' and 'a lot' in Figure 3(a)) fare better aided by Light Bulb than Tick/Cross. Nevertheless, Light Bulb leads to the lowest grades among children who were certain

that they had found the right information to respond to the search prompt (Figure 3(b)). Another interesting pattern emerging from Figure 3(b) is that children who stated not finding the information they were seeking still scored around 90 points using Light Bulb, decreasing to closer to 80 on Tick/Cross and No Emoji. These findings would indicate that besides scaffolding in the form of enriched interfaces based on self-efficacy assessment, children require clear instructions that can help them navigate SERP to locate relevant resources and within them the information they need to complete their assignments.

Children who enjoy searching are also more likely to explore SERP resources, regardless of the interface used. On the other hand, children who do not enjoy searching keep a traditional linear exploration strategy and select top-ranked resources when using No Emoji; they do explore farther on the SERP with the help of emoji-enriched interfaces (Figure 4(a)). Children who believe they have not found all the information they were looking for were willing to click on lower-ranked SERP resources only when interacting with emoji-enriched SERP, selecting top-resources instead when using No Emoji. On the other end of the spectrum, children who confidently asserted that they have found what they were seeking clicked on resources positioned 4th in the ranking, regardless of the interface used (Figure 4(b)). It seems that children who do
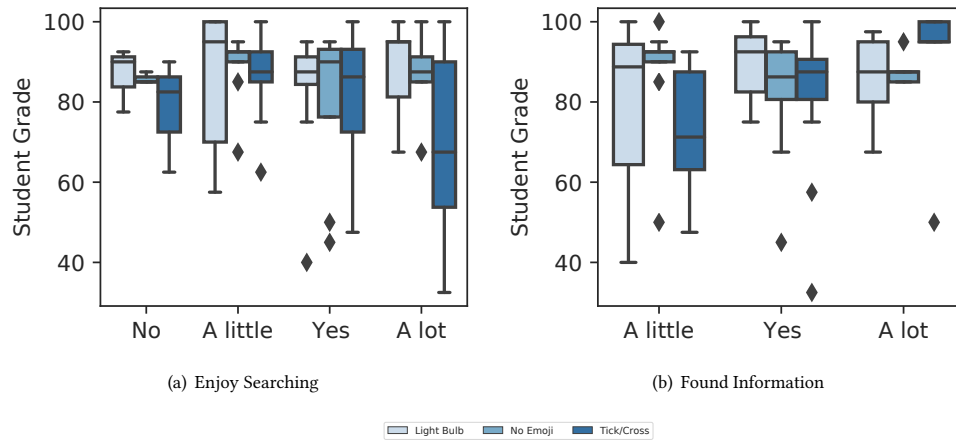
(a) Enjoy Searching

(b) Found Information

Light Bulb   No Emoji   Tick/Cross

Figure 3: Children's perception on efficacy for searching and task completion w.r.t. search success.

(a) Enjoy Searching

(b) Found Information

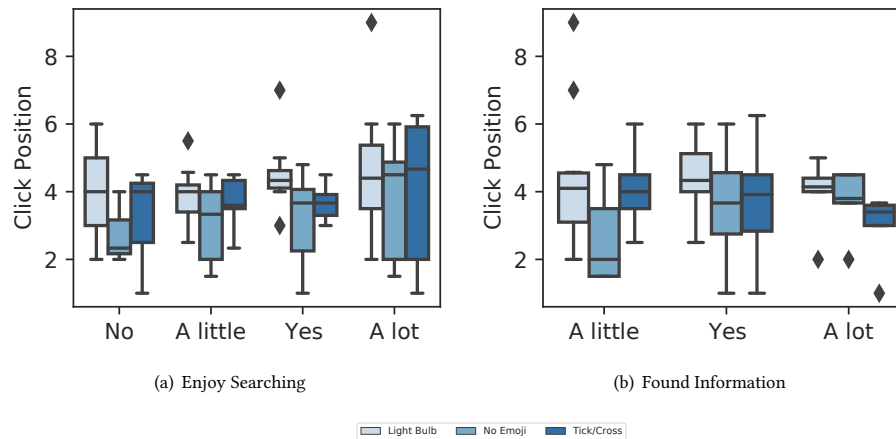Light Bulb   No Emoji   Tick/Cross

Figure 4: Perception on efficacy for searching and task completion w.r.t. ranking position of resources children interact with.

not particularly enjoy searching and are less confident in their ability to extract information from retrieved results benefit from emoji-enriched interfaces.

Do search behavioural traits impact search success? To answer RQ3, we look for common traits among children with similar search success. Is it possible to recognise specific traits common to children with similar success rates? If so, what is the impact of the enriched SERP on their performance? In the end, does the use of enriched SERP influence in a positive way overall search success?

As illustrated in Figure,5(a) children with Highsearch success benefit from enriched interfaces—search success scores are high regardless of the interface used; still, the scores are slightly higher using enriched ones. Children with Low search success generally obtained low grades regardless of the interface used for seeking purposes. We attribute this to them being disengaged with the task. For example, using Tick/Cross, children with Low search success clicked on fewer results and spent the least amount of time

on their search task (shown in Figure 5(b) and Figure 5(f), resp.). Even more so, we attribute these results to the inability of children with Lowsearch success to extract the right information from the retrieved results. In other words, this user group clicked on more known relevant resources than children with Medium or Highsearch success; even so, they were not capable of using these resources to answer the pertinent questions on Ancient Rome (see Figures 5(d) and 5(e)). This would suggest that the most effective scaffolding for this group of children would be some form of specific guidance to help them understand the different steps of the searching online process. We could envision the design of an intelligent search companion to support children and provide them with the initial necessary scaffolding [13, 15]. Figure 5(a) tells an interesting story about children with Medium search success, who did not get much support from enriched interfaces; if anything, enriched interfaces hurt their overall grade. We surmise that this could be due to children in this group finding emojis confusing or distracting. The
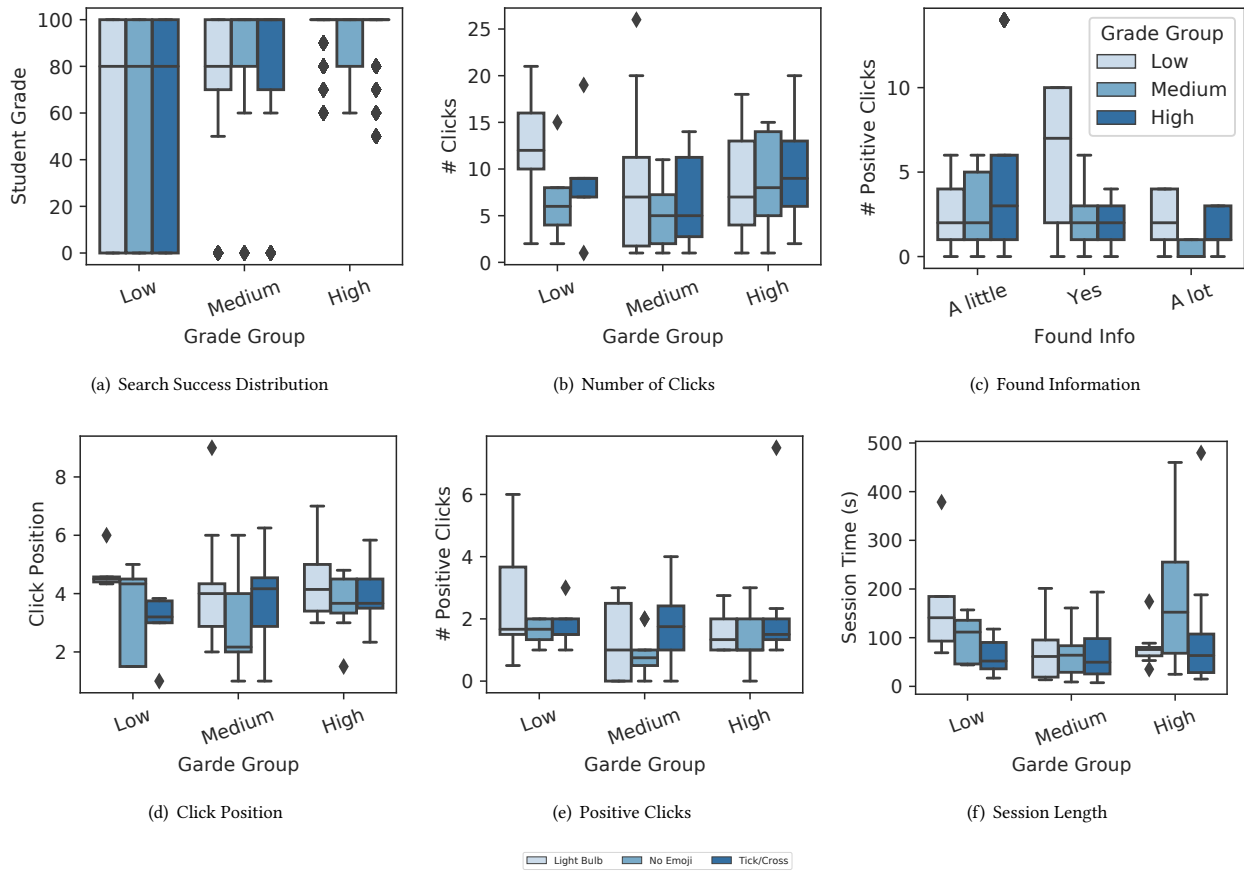
(a) Search Success Distribution

(b) Number of Clicks

(c) Found Information

(d) Click Position

(e) Positive Clicks

(f) Session Length

Figure 5: Exploration of search behavioural traits based on search success.

confusion is notable in Figures 5(b) and 5(e): children in this group had the highest amount of clicks using Light Bulb, yet the least using Tick/Cross; paradoxically, they clicked more frequently on known relevant results using Tick/Cross. Thus, we should design to increase motivation and avoid distraction in this user group.

Trends emerging from Figure 5 evidence that creating user models that capture children's search behavioural traits and mapping these models to the type of interface support these users need is a complex endeavor, one requiring deeper examination beyond search success.

Limitations. The size of the user sample could be perceived as a limitation, even if it is rare to find in literature larger studies involving this type of population, as working with children in a classroom has many constraints and ethical related restrictions. Nonetheless, as we collected data in a naturalistic study, we were able to run an in-depth intensive analysis, providing valuable insights. It is also worth noticing that unlike other studies related to searching in the classroom and therefore involving children, the study we refer to here was run completely online due to the COVID-19 pandemic—this increased cognitive overload for children.

4 CONCLUSIONS AND FUTURE WORK

We have described our analysis of user data to guide innovative SERP design to aid children in noticing the (most) relevant/useful results when performing classroom-related search tasks. In order to better understand children's behaviour, we looked for links across three data types: implicit—search logs, explicit—children's perception of different facets of their overall search experience, and expert assessments—produced by educators when grading the answers provided at the end of the task. In this, we were guided by the FHAB lenses, introduced by Aliannejadi et al. [1].

As mentioned by Bilal [3], "children are *unique* young users and not simply *short adults*". Our results further supported this statement. Even though study participants have the same age and belong to the same school grade, we were not able to identify a particular SERP that equally benefited, i.e., supported, all the children in the study. Dividing them into 3 groups based on expert assessments, we could see some interesting trends: we could recognise the best type of scaffolding working for children in the High search success group and how children in the Low search success group need some form of extra guidance to make sense of the searching process and its complexity. Still, we could not definitively capture what was going on with the Medium search success group. Thus, we reckon

there is a need to explore a multi-faceted space where children can play different roles and so have different needs and preferences when searching. We suppose that children in the `Medium` search success group have not benefited from the enriched SERP because they were non-motivated and/or distracted searchers as described by Druin et al. [5]. For this reason, different types of scaffolding solutions should be studied to respond better and adapt to children's needs. For instance, we suggest making inquiry tasks more engaging by including emotional elements as described [14] so as to entice non-motivated searchers. We also recommend avoiding presenting additional attractive material not strictly pertinent to the search in order to target distracted searchers. Overall, our aim remains to provide adaptive SERP that support children while developing their searching skills from beginners to independent searchers and so scaffold the learning process in the right way and at the right time.

REFERENCES

[1] Mohammad Aliannejadi, Monica Landoni, Theo Huibers, Emiliana Murgia, and Maria Soledad Pera. 2021. Children's Perspective on How Emojis Help Them to Recognise Relevant Results: Do Actions Speak Louder Than Words?. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*. 301–305.

[2] Oghenemaro Anuyah, Ashlee Milton, Michael Green, and Maria Soledad Pera. 2019. An empirical analysis of search engines' response to web search queries associated with the classroom setting. *Aslib Journal of Information Management* (2019).

[3] Dania Bilal. 2010. The mediated information needs of children on the Autism Spectrum Disorder (ASD). In *Proceedings of the 31st ACM SIGIR Workshop on Accessible Search Systems, Geneva, Switzerland*. Citeseer, 42–49.

[4] Arthur Câmara, Nirmal Roy, David Maxwell, and Claudia Hauff. 2021. Searching to learn with instructional scaffolding. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*. 209–218.

[5] Allison Druin, Elizabeth Foss, Hilary Hutchinson, Evan Golub, and Leshell Hatley. 2010. Children's roles using keyword search interfaces at home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 413–422.

[6] Carsten Eickhoff, Pieter Dekker, and Arjen P De Vries. 2012. Supporting children's web search in school environments. In *Proceedings of the 4th Information Interaction in Context Symposium*. 129–137.

[7] Jerry Alan Fails, Maria Soledad Pera, Oghenemaro Anuyah, Casey Kennington, Katherine Landau Wright, and William Bigirimana. 2019. Query formulation assistance for kids: What is available, when to help & what kids want. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. 109–120.

[8] Tatiana Gossen, Juliane Höbel, and Andreas Nürnberger. 2014. Usability and perception of young users and adults on targeted web search engines. In *Proceedings of the 5th Information Interaction in Context Symposium*. 18–27.

[9] Jacek Gwizdka and Dania Bilal. 2017. Analysis of children's queries and click behavior on ranked results and their thought processes in google search. In *Proceedings of the 2017 conference on conference human information interaction and retrieval*. 377–380.

[10] Alexander Halavais. 2017. *Search engine society.* John Wiley & Sons.

[11] Michel Jansen, Wim Bos, Paul van der Vet, Theo Huibers, and Djoerd Hiemstra. 2010. TeddIR: tangible information retrieval for children. In *Proceedings of the 9th international conference on interaction design and children*. 282–285.

[12] Yvonne Kammerer and Maja Bohnacker. 2012. Children's web search with Google: the effectiveness of natural language queries. In *proceedings of the 11th International Conference on Interaction Design and Children*. 184–187.

[13] Monica Landoni, Emiliana Murgia, Theo Huibers, and Maria Soledad Pera. 2020. You've Got a Friend in Me: Children and Search Agents. In *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. 89–94.

[14] Monica Landoni, Maria Soledad Pera, Emiliana Murgia, and Theo Huibers. 2020. Inside Out: Exploring the Emotional Side of Search Engines in the Classroom. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. 136–144.

[15] Silvia B Lovato, Anne Marie Piper, and Ellen A Wartella. 2019. Hey Google, do unicorns exist? Conversational agents as a path to answers to children's questions. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. 301–313.

[16] Neil Mercer. 1994. Neo-Vygotskian theory and classroom education. *Language, literacy and learning in educational practice* (1994), 92–110.

[17] Ian Rowlands, David Nicholas, Peter Williams, Paul Huntington, Maggie Fieldhouse, Barrie Gunter, Richard Withey, Hamid R Jamali, Tom Dobrowolski, and Carol Tenopir. 2008. The Google generation: the information behaviour of the researcher of the future. In *Aslib proceedings*. Emerald Group Publishing Limited.

[18] Nicola Yelland and Jennifer Masters. 2007. Rethinking scaffolding in the information age. *Computers & Education* 48, 3 (2007), 362–382.