2020

Say It with Emojis: Co-Designing Relevance Cues for Searching in the Classroom

Mohammad Aliannejadi
University of Amsterdam

Monica Landoni
Università della Svizzera Italiana

Theo Huibers
University of Twente

Emiliana Murgia
Università degli Studi di Milano-Bicocca

Maria Soledad Pera
Boise State University

1st Joint Conference of the Information Retrieval Communities in Europe, CIRCLE 2020, July 6, 2020 - July 9, 2020
Say it with Emojis: Co-designing Relevance Cues for Searching in the Classroom

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

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

Theo Huibers
t.w.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, Idaho

ABSTRACT
Search Engine Result Pages (SERP) include snippets of retrieved resources as a means to help searchers select the ones that satisfy their information needs. This way, result relevance can be determined by scanning through snippets, an exercise that requires experience with reading, understanding, and assessing the value of a document. These skills that primary school children are still developing and thus are not yet proficient with. As web search tools are essential to support children learning at school and home, we explore how to help young searchers in making informed relevance assessments while conducting searches in a classroom. In this paper, we describe a collaborative design exercise involving primary school children as co-designers: we asked them to examine interfaces with combinations of different emojis to help them assess the usefulness of results in SERP—a crucial factor to determine relevance for the classroom. This activity made our child experts engage with the design exercise while enabling us to collect their judgments so as to get a better sense of the user requirements for this age group. Here we discuss the main design issues emerging from the analysis of children’s preferences, the rationale behind them, comments and concerns raised, and alternative proposals children sketched.

1 INTRODUCTION
Children’s search behaviors have been well-documented [6–8, 15, 17, 35]. Among the main ones, we find children’s struggles with query formulation – from inability to create succinct queries to favoring the use of natural language questions. Another issue that emerged is children’s linear exploration of Search Engine Result Pages (SERP) and their tendency to favor results positioned higher in the SERP, regardless of the relevance of these results. Researchers have attempted to ease query formulation for children, both from interface and algorithmic perspectives [15, 27]. Yet, we see seldom efforts dedicated to addressing the gaps related to children and their judgment of results presented on SERP in response to inquiries submitted to popular or kid-friendly search engines (SE), like Google or Kiddle, respectively.

SERP snippets often include the title, source, and sample text from a given resource. While finding it easy to scan through the brief snapshot and determine the relevance of the corresponding resource, for young searchers, this task is not as straightforward. As stated by Gossen et al. [18], children’s exploration of SERP tends to be more visual, i.e., influenced by highlighted terms. In other words, children exhibit a “cued visual jump” strategy [18], one that makes them jump sequentially to the following resource after reading the highlighted terms in a given snippet. Moreover, reports in [18] indicate that children pay attention to thumbnails and other media embedded in snippets. This is one of the reasons why we posit that icons could serve as visual cues that inform the relevance of resources listed in a SERP.

We are interested in exploring whether visual cues can influence children’s selection of useful resources in response to their queries, regardless of the position of such a result in the SERP. To control scope, and ease comparison and contextualization with respect to the state-of-the-art, we adopt the framework defined in [21], which establishes four pillars for design and evaluation of information retrieval systems for children: (i) strategy, (ii) user group, (iii) task, and (iv) context. In our case, we define (i) SERP display, (ii) children in primary five, (iii) online inquiries pertaining to topics common among primary four curricula, and (iv) classroom setting. In this initial iteration, we focus in fostering children’s selection of results that are of an educational nature, i.e., aligned with the context of the information-seeking tasks under study. Consequently, we use emojis to signal classroom-aligned vs. more general sources. The choice of emojis is driven by existing research outlining children’s ability to relate to emojis [5, 16, 26, 32]. When children are the audience under study, emojis have been shown to be effective as a means for children to experience preference, which is why it stands to reason that they will consider them a natural clue for relevance, which is the hypothesis of our work. There is not, however, research related to which emojis would better serve as clues. Consequently, we start our exploration by conducting a study in which we use children as expert co-designers to (i) better understand which emojis are more intuitive to understand, in terms to serving as a relevance clue, and (ii) identify requirements from children as to what constitute a better emoji that can foster completion of successful searches conducted in the classroom.

Insights from co-designing activities serve as a foundation for research in human-computer interaction and information retrieval, in terms of fostering design of interfaces and algorithms that provide the scaffolding needed to support children’s search in the classroom while learning how to conduct inquiry tasks effectively.

*Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).*
2 RELATED WORK

We anchor our work in two main areas: SERP design and co-design with children. We briefly discuss relevant works below.

Gwizdka and Bilal [19] examine children’s interactions with Google. The authors found that when exploring SERP, younger children (age 11) are less deliberate on their choice of results to click, as they tend to favor top two-ranked results in a SERP. Amuyah et al. [3] describe the type of resources displayed on SERP in response to children’s inquiries in the classroom, and argue for the need to prioritize classroom-related materials in SERP, as well as aid children in identifying the most suitable results for classroom use. Lurie and Mustafaraj [25] also highlight the disparity in children’s opinions on Knowledge Panels (which provide contextual information about news) offered on SERP to gauge the credibility of online news. Similarly, Abdullah and Basar [1] study young searchers’ behavior when scanning SERP to identify trustworthy resources. Their findings reveal that children seldom consider source reputation when determining its relevance. As an attempt to offer users a richer SERP, Demartini and Siersdorfer [11] include sentiment-related information along with snippets. They propose the use of green and red colors to flag the positive and negative connotation of content listed on a SERP. This solution, however, aligns with enhancing SERP design for adult searchers, not children.

The literature pertaining to children as co-designers is rich, starting from early work by Druin [13], who describes the four main roles children can play in design: user, tester, informant, and design partner. The most widely adopted approaches for engaging children as co-designers are participatory design by Read et al. [30] and cooperative inquiry by Druin [12]. Crucial is also to find ways to give children’s opinion the right importance and consideration [20, 29]. We see from the literature that children have mostly been involved in the ideation phase of design [22]. Instead, they have seldom actively contributed to other phases of design, e.g. preparing sketches and storyboards and acting as experts in early usability evaluations. Examples of such contribution come from van Doorn et al. [33, 34], who explored how children could collaborate in many different ways, from interviewers to full co-researchers. They argue that children uniquely have the necessary knowledge to engage and understand their peers. On similar lines, Salian et al. [31] debate on how children could act as heuristic evaluators.

In our investigation, we reiterate the aforementioned need for richer SERP design that can ease the choice of useful results for children in the classroom. We attempt to address such a concern by conducting an exploratory quest guided by children’s expertise to identify emojis that can serve as relevance cues.

3 METHOD & DATA

We describe in this section the participants, data, and protocol, we considered in our exploration.

Co-Designers. We recruited as co-designers (i.e., study participants) 9 children (ages 10 to 11) from a primary five classroom in Italy. To align with study goal, we sought children who (i) have regular exposure and instruction related to search tools and (ii) have participated of previous studies involving co-designing interfaces for information retrieval systems tailored to children and the classroom [21, 28]. It is worth mentioning that recruitment was on voluntary basis. We obtained parental consent in advanced; the study was approved by the local Ethics committee.

Mock-up Interfaces. In collaboration with experts in education, we prepared 3 mock-up interfaces, as shown in Figure 1, which we used to stimulate discussion with children. For SERP-generation, we used Bing API [4], with safe search functionality enabled. Each mock-up includes emojis to cue results that are useful (or not) for the classroom; neutral results were not associated with any emoji. Results for each query were labeled by an expert educator. Note that informed by the lessons learned from [9, 21], children often associate result relevance with its usefulness to complete search tasks aligned with school curriculum. Thus, in our study, we treat relevant and useful for the classroom as equivalent.

(1) **Mock-up 1** (Figure 1a) has **thumb-up and down** icons next to relevant and irrelevant results, in order to attract children’s attention via the use of a very popular icon used to express ‘I like it’ in social media.

(2) **Mock-up 2** (Figure 1b) displays two fun icons: **rainbow and poo** to hint of (non-)relevance of the associated results.

(3) **Mock-up 3** (Figure 1c), showcases the classic **traffic light** icon.

In designing the mock-ups, we exploited different metaphors as a conduit to prompt discussion. Consequently, we paid particular attention to offering children a stimulating variety of options that...
We include in Table 2 a snapshot of its the short version, which we also invited to sketch a possible alternative. We include the full adaptation to better suit the age of the target population under study. We used two different tools to stimulate and gather feedback.

(1) In **Phase 1**, we adopted an existing interview scheme, which we used as an early inspection method involving experts. Specifically, we leveraged the Cognitive Walkthrough approach [10], where children acted as experts of other children’s needs and preferences when searching in the classroom. In this instance, while walking children through a search activity, we asked not only their individual preferences but also which of the mock-ups they reckoned their classmates would choose. We also asked children to consider what would happen when searching for other school-related subjects. At the end of the search walkthrough, we encouraged the children to comment on what they liked and disliked the proposed mock-ups. We also invited to sketch a possible alternative. We include the full interview questionnaire in Table 1.

We also used a standard User Experience Questionnaire [2, 23]. We include in Table 2 a snapshot of its the short version, which we adapted to better suit the age of the target population under study. We did so by simply adding a note to clarify terms deemed too difficult for children, e.g., obstructive, and providing more examples of how to express the assessment of each of the listed terms.

### Table 1: Student interview scheme, translated from Italian, the language spoken by participants.

<table>
<thead>
<tr>
<th>ID</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Which of the 3 interfaces would you use if you were looking for information on the tornado topic? Why?</td>
</tr>
<tr>
<td>Q2</td>
<td>Which of the 3 interfaces do you think kids your age would use if they were looking for information about tornadoes? Why?</td>
</tr>
<tr>
<td>Q3</td>
<td>Which of the 3 interfaces would you use if you were looking for information on another topic for school research? Why?</td>
</tr>
<tr>
<td>Q4</td>
<td>Which of the 3 interfaces do you think kids your age would use if they were looking for information on other topics for school research? Why?</td>
</tr>
<tr>
<td>Q5</td>
<td>Please add your comments on what you like or dislike in the 3 interfaces. What changes would you propose?</td>
</tr>
<tr>
<td>Q6</td>
<td>If you had to design a new interface to use for school research material, what would it look like?</td>
</tr>
<tr>
<td>Q7</td>
<td>You can add a drawing and an explanation of why you would do it that way.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ID</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q8</td>
<td>Which of the 3 interfaces do you think kids your age would use if they were looking for information about events? Why?</td>
</tr>
<tr>
<td>Q9</td>
<td>Which of the 3 interfaces would you use if you were looking for information on another topic for school research? Why?</td>
</tr>
<tr>
<td>Q10</td>
<td>Please add your comments on what you like or dislike in the 3 interfaces. What changes would you propose?</td>
</tr>
<tr>
<td>Q11</td>
<td>If you had to design a new interface to use for school research material, what would it look like?</td>
</tr>
<tr>
<td>Q12</td>
<td>You can add a drawing and an explanation of why you would do it that way.</td>
</tr>
</tbody>
</table>

**Table 2: User Experience Questionnaire for each mock-up.**

<table>
<thead>
<tr>
<th>Likert scale</th>
<th>Supportive</th>
<th>Easy</th>
<th>Efficient</th>
<th>Clear</th>
<th>Exiting</th>
<th>Interesting</th>
<th>Inventive</th>
<th>Leading Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obstructive</td>
<td>□ □ □ □ □</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complicated</td>
<td>□ □ □ □ □</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inefficient</td>
<td>□ □ □ □ □</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusing</td>
<td>□ □ □ □ □</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boring</td>
<td>□ □ □ □ □</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Interesting</td>
<td>□ □ □ □ □</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td>□ □ □ □ □</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usual</td>
<td>□ □ □ □ □</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

at the same time (i) were familiar to them, (ii) suitably represented a dichotomy for relevance judgment, and (iii) showcased different combinations of colors and shapes.

**Tools to Stimulate and Gather Feedback.** We used two different tools to stimulate children and engage them in co-design activity. First, we adopted an existing interview scheme, which we used as an early inspection method involving experts. Specifically, we leveraged the Cognitive Walkthrough approach [10], where children acted as experts of other children’s needs and preferences when searching in the classroom. In this instance, while walking children through a search activity, we asked not only their individual preferences but also which of the mock-ups they reckoned their classmates would choose. We also asked children to consider what would happen when searching for other school-related subjects. At the end of the search walkthrough, we encouraged the children to comment on what they liked and disliked the proposed mock-ups. We also invited to sketch a possible alternative. We include the full interview questionnaire in Table 1.

We also used a standard User Experience Questionnaire [2, 23]. We include in Table 2 a snapshot of its the short version, which we adapted to better suit the age of the target population under study. We did so by simply adding a note to clarify terms deemed too difficult for children, e.g., obstructive, and providing more examples of how to express the assessment of each of the listed terms.

**Protocol.** We conducted co-design sessions online, as schools had to close for a medical/health emergency. We organized these sessions in 3 phases with one researcher acting as facilitator.

(1) In **Phase 1**, the researcher presented children the three mock-up interfaces (see Figure 1), each including a SERP for the same given query. For the topic of the query included in the mock-up interfaces, we followed the framework and protocol in [21] and settled on common subjects in the 4th and 5th grades: tornadoes. In each mock-up, we used a different emoji to highlight the usefulness (or lack thereof) of the corresponding result. The researcher kept children engaged by asking them to imagine how useful other children would find the three proposed interfaces. Also, we reminded the students that they were acting as experts, knowing what would better work, and be useful for other children like them. We stressed the fact that they knew this area better than any adults, including their teachers.

(2) In **Phase 2**, still online, the researcher, who plays the role of facilitator, guided children so that they could frame the applicability of the mock-ups in the context of searching in the classroom. It was aimed initially for one specific task and then for generic ones, moving the focus from their perspective to how their peers would react. Children were invited to act as assessors and discuss preferences and reasons behind them. At the end of the discussion, children were asked to pick their most liked and disliked features across the three interfaces. Then, for each interface, they were invited to fill a short questionnaire (see Table 2). This questionnaire enabled us to gather insights regarding desirable features for emojis that would better help children understand the usefulness of resources retrieved in response to online inquiries related to curriculum topics.

(3) In **Phase 3**, children, assumed the creative role of designers and worked off the line on sketches to represent new interfaces and icons for the same purpose. Here, they were advised to draw the sketches using personal devices of their choice, such as tablets, smartphones, and PCs. Upon completion of the three phases, children shared with researchers via email their responses for the interview and questionnaires, along with the sketches created during Phase 3.

**Data.** Children’s preferences and the reasons behind them provided us with essential insights. For analysis purposes, we considered children’s responses to the questionnaires in Tables 1 and 2, as well as their suggestions in the form of sketches and a brief description of suitable emojis. Observations from teachers who administered the protocol, as mentioned earlier, were also valuable in better understanding the outcome of this design exercise.

### 4  ANALYSIS & DISCUSSION

In this study, we intended mock-ups to be used as probes to engage our co-design experts and stimulate their creativity. Therefore, our findings are meant to contribute towards better understanding of young searchers’ needs and to provide necessary grounds for the design of SERP that can support children in searches in the school context. In the rest of this section, we discuss the main issues and discoveries emerging from this co-design effort.

**Inclusiveness.** From the discussion with children and answers to questionnaires, it is clear that inclusiveness is an important
feature when choosing emojis for enhancing the SERP. Children commented on the use of red and green as not suitable for color-blind readers (mock-up 3). One child, in particular, peculiarly used the questionnaire to express this concept. When asked to put an X in the row with Confused and Clear, he put two: one next to confused for people who would not be able to distinguish green and red, and one next to clear for people who would see the difference. This shows how a standard tool for experience evaluation can be used in a flexible and meaningful way when running design activities. This interaction also showcased how the child was acting as an expert and providing an answer based on other children he knew of.

Fun vs. Formal. Mock-up 2 was judged as the most fun, but at the same time, children commented on it not being suitable to be used in class, “What do you think my teacher would think of it?” was one of such comments. It was interesting to see how answers to questionnaires served to clarify the answers given during the interactive interview session. When directly asked about what other children would prefer, mock-up 2 had the most votes. Yet, when reflecting on why, children mostly highlighted how silly it was; further, in the questionnaires, children gave it the lowest score when reflecting on why, children mostly highlighted how silly it was one of such comments. It was interesting to see how answers vs. adult results (see Figure 2b). In this particular case, the child line with this idea is the use of a divider emojis to be easy to recognize when scrolling while keeping up independently, suggested the same creative new design, i.e., a effective also while scanning the SERP quickly. Two of the children, in combination with an easy-to-spot difference in color to be ef-
up and down binary change: between them even on the run. The use of a simple metaphor of scrolling quickly in a page of results. From this, it arises that icons described it. However, children expressed their concerns about not was how one of our co-design experts de-
“hands save the day” Mock-up 1 was judged to be easy and clear; Easy to Recognize. Mock-up 1 was judged to be easy and clear; “hands save the day” was how one of our co-design experts described it. However, children expressed their concerns about not being able to distinguish the two icons, thumbs up and down, while scrolling quickly in a page of results. From this, it arises that icons should have clear and recognizable features to let users discriminate between them even on the run. The use of a simple metaphor of binary change: up and down or on and off have to be reinforced in combination with an easy-to-spot difference in color to be ef-fective also while scanning the SERP quickly. Two of the children, independently, suggested the same creative new design, i.e., a bulb switched on and off bulb, which was motivated by the need for emojis to be easy to recognize when scrolling while keeping up with the inclusiveness requirements. Another proposed emoji in-line with this idea is the use of a divider to isolate child-friendly vs. adult results (see Figure 2b). In this particular case, the child stated emphatically that there was no need to cue for useful results. This is because she expected the SE to present only relevant results, hence the only hint needed was to find child-friendly ones among the results in the SERP.

Remarks. Our study has elicited user requirements to guide the next stage of production design for innovative SERP. It has also provided an insight into how children interpret the concept of relevance when searching for school-related tasks. In particular, children judge as “good” those results that are useful, reliable, and trustworthy. Literature confirms how children naturally trust technology and struggle to assess the quality of search results critically [14]. Thus, it is important to design search tools that foster the development of such skills by providing extra hints in a clear form, such as emojis.

Limitations. Our study and findings also have some limitations that we are aware of and plan to address in the future. Recent developments of COVID-19 caused the closure of schools. Thus, we conducted the studies via online tools such as Skype. As a result, children could have been distracted or influenced by their environment, being at home. Also, the manner in which the co-design sessions were conducted, could have affected the children’s creativity, as well as their interaction with the researcher. Currently, we only use binary relevance values. In future iterations of our work, we will explore more nuanced methods of assessing their interaction with the researcher.

5 CONCLUSIONS & FUTURE WORK

We have explored how to design emoji-enhanced SERP, as means of visual cues to bias selection of resources that are not only relevant to children’s information needs but also the context of the search. We conducted a number of collaborative design sessions with children in primary five, who were offered three mock-up interfaces to stimulate their critical thinking and creativity.

Results emerging from this initial iteration of our work reveal a need for further exploration on how to generate and select what other cues should be considered, in order to help children locate relevant resources, when using their preferred SE—often mainstreams ones, like Google, which were not designed with kids in mind. Issues emerging from this sessions will drive the design of innovative SERP to provide scaffolding to children searching at school. We will start by producing a series of disposable prototypes to be evaluated by a larger cohort of children of similar age. We will pay particular attention to including a representative sample of children who are technically savvy vs. novices, as this could bring in different needs and preferences as in studies with adults [24]. From co-design experts’ feedback, it seems that emojis could also serve as cues for different types of resources included in SERP (magazines for children, video, Wikipedia, etc.). This is another research path that needs consideration. Adding emojis to a SERP introduces a new kind of bias towards the results that the system favors. That is why in the future, we plan to run a comparative user study where we will ask children to complete some search tasks and provide them with regular SERP, as well as emoji-equipped SERP. We expect to study the effect of emojis on helping students find useful information. Also, we plan to investigate if emoji bias is stronger than order bias in a SERP.
ACKNOWLEDGMENTS
We would like to thank the co-design experts who were willing to join online to share their feedback. We also appreciate the work by Mihail Kicev, who designed the mock-ups used in our co-design exercise.

REFERENCES