

Children’s Perspective on How Emojis Help Them to Recognise Relevant Results: Do Actions Speak Louder Than Words?

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.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, Idaho

ABSTRACT

We discuss the exploratory study we conducted to better understand children’s ability to recognise relevant results when searching in the classroom. Teachers in two European schools sharing the same language assigned their students (ages 10 and 11) an online information discovery exercise about a history topic covered in class. For this, children used a classic search interface and two novel ones enriched with emojis associated to relevant vs. irrelevant results. At the end of the exercise, children filled out a post-task questionnaire meant to elicit their perception on usability of the interfaces. Guided by four lenses, we analyse our findings and discuss whether (i) emoji-enriched interfaces lead to better performance for children using a search engine in the classroom and (ii) “actions speak louder than words” when looking at children’s search experience. We learned various lessons from our examination of children’s search behaviour that will guide the design of future interfaces, including the fact that emoji-enriched interfaces result in significant improvement in terms of children identifying relevant resources.

ACM Reference Format:

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 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '21), March 14–19, 2021, Canberra, ACT, Australia*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3406522.3446058>

1 INTRODUCTION

Despite the advances in Information Retrieval, children continue to struggle with effectively using search tools to locate relevant information [6]. Search literacy instruction could afford children the know-how to best use existing search tools [11], but this is a

*Co-founder of WizeNoze.

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 ACM 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](https://permissions.acm.org).

CHIIR '21, March 14–19, 2021, Canberra, ACT, Australia

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-8055-3/21/03...\$15.00

<https://doi.org/10.1145/3406522.3446058>

long term solution. Voice-controlled digital assistants could help bypass literacy-based barriers for children seeking for information [6]. Unfortunately, they are not yet at a stage to respond to cognitive and developmental needs of children [13, 16, 18].

As an attempt to ease the information seeking process, researchers and industry practitioners have designed tools tailored for children [8], some even explicitly supporting completion of school-related assignments. Notable examples include tools to aid query formulation [15], in addition to Infotopia.info, an academic search tool based on curated resources for students and teachers [9] and webfor-classrooms.com, a search tool designed to enable students to locate online content that matches their needs and reading skills [17]. Still, children often turn to popular search engines (SE) like Google or Bing, which have been designed for adults [15]. In their seminal work, Jochmann-Mannak et al. [10] report on lessons learned from their study comparing performance on children’s interfaces vs. Google. They emphasise the need to enhance interfaces of traditional SE by adding pictures or categories to search results. Aliannejadi et al. [1] follow this mandate in their co-design study in which children—treated as experts—design icons that could enrich search engine result pages (SERP), in terms of offering relevance cues. Outcomes from their work, however, indicate that children often design icons that would help *their peers*, but do not necessarily feel these icons would help themselves. This leads us to explore whether what children like and want aligns with what they need and use.

A wide range of scholarly works focus on the difference between what users prefer (perception) and what they do (action). Liu et al. [14] discuss the inconsistencies between perception on usefulness of clicked SERP resources with respect to the content of the resources themselves, as from study results it was clear that “the search was not successful”. When comparing spoken language vs. traditional keyword-based interfaces, results of studies with children and adults demonstrate that while users favour spoken interfaces, in practice, search sessions via spoken interfaces produce fewer clicks, longer queries, and longer overall session length [13, 19]. With children as the primary stakeholders, discrepancies between perceptions and actions are even more apparent. Zhao et al. [20] uncovered that while children claim to understand the concept of privacy, they would not always take effective action to address a related risk. Danovitch [6] showed that children’s access and engagement with search tools rarely translates to them making effective use of the tools to discover information.



Figure 1: Three mock-ups interfaces enriched with emojis as cues for (non-)relevance.

We aim to set the groundwork towards understanding the aspects that should steer the design of useful and usable search technology for children to use in an **educational context**. In particular, we explore the characteristics that directly impact perceived vs. actual user experience (**UX**) with **interfaces** that can ease children's online information seeking in the classroom setting. UX considers aesthetic and emotional elements together with usability varied factors from how easy and pleasant a tool is to use to the level of satisfaction its performance gives to users. To guide our exploration, we outline four lenses grounded on UX factors, which we refer to as **FHAB lenses** and in the context of this work are defined as:

- **Favourite**: Which interface do the children like most? It is imperative that children like the tools they use in an educational context.
- **Helpful**: Which interface would the children prefer to use in the future to perform an information discovery tasks related to primary school curriculum?
- **Advised**: Which interface would the children suggest to their school friends to perform a search task?
- **Best Performing**: Which interface is most useful and led the children to be the most accurate, i.e. click on results relevant to their inquiry?

These lenses are based on findings from previous studies [1, 12, 13], enabling us to devise the most suitable factors to describe children's search experience in the classroom. As aesthetic proved to play an important role in the interaction between SE and children, we look at the interface children *favoured* most. We consider *helpfulness* as a usability measure to account for how satisfying children believe the experience will be (to avoid the immediacy of the novelty effect we look at SE interactions mediated by a close future projection). Via *advised*, we probe how children feel about recommending a particular interface to their peers in the classroom context. To examine children's *performance*, we include a system-related measure by observing how accurate their clicks of relevant resources are.

Following the four pillars established in [13] to guide the design and evaluation of information retrieval systems for children, we scope our exploration for: (i) children in primary five (ages 10 to 11), (ii) looking for online resources to (iii) complete classroom-related inquiry tasks using (iv) SE with(-out) enhanced interfaces. Specifically, we invited 100 children to partake in an information discovery assignment related to a common classroom subject in primary five: Ancient Rome. By focusing on a common classroom-related task (co-designed with teacher experts), we mitigate the bias of artificial search tasks that often drive research works in this area [5]. To complete the study, children were presented with 3 different mock-up interfaces as shown in Figure 1: a traditional one, one enhanced with emojis designed by children to signal (non-)relevance and one enhanced with common emojis. This was followed by a short

post-task questionnaire, in which we inquired about children's perceived support received from the different interfaces to complete the assigned search tasks. Analysis of post-task questionnaire responses and generated query logs based on FHAB lenses allow us to answer: **RQ1**. To what extent do **emoji-enriched interfaces** lead to better performance for children using SE in the classroom? and **RQ2**. Do **actions speak louder than words** regarding interface used to complete school-related assignments?

In the case of RQ2, we investigate if children liking something means they will see it as ideal for future use? Will they advise their classmates to use the same? Are favoured interfaces the ones that lead to the best performance? We hypothesised that the most logical answer to many of these questions would be "yes". However, preliminary results show that this is not be the case, suggesting the need to pay even more attention to the actual behaviour and spoken explanations of children when designing SE and other related tools for the classroom. Our findings impact the topic of child SE: children's advice and feedback on the design is always important but should be taken into account in a weighted way.

2 STRUCTURE OF THE STUDY

In this section, we describe how we carried out the research, with how many and which participants, what data and what protocols.

Participants. Upon agreement with schools' principals and teachers, we recruited 100 Italian-speaking children (ages 10 to 11) from primary five classrooms in Italy and Switzerland. Participants have varying levels of exposure and instruction related to search tools, making them representative of primary five populations.

Interfaces & Labels. Inspired by Aliannejadi et al. [1], who explore the use of emojis to help children identify relevant resources when seeking information online, we propose three mock-up interfaces (Figure 1). For SERP generation, we used Bing API [2] (in Italian), with safe search enabled. Expert educators then labelled each of the top 10 results retrieved in response to the search prompts in our protocol as *relevant*, *non-relevant*, and *neutral*. As children often associate result relevance with its usefulness to complete a task [4, 13], we treat relevant and useful for the classroom as equivalent.

- (1) **Mock-up Light Bulb** has *light bulb on* and *off* icons next to relevant and non-relevant results; neutral results were not associated with any emoji. This mock-up was inspired by emojis designed by children to support the search process [1].
- (2) **Mock-up Tick/Cross** displays traditional *check-marks* and *crosses* to indicate the (non-)relevance of results.
- (3) **Mock-up No Emoji**, simulates a classic SE interface with no emojis. This mock-up serves as "control".

Table 1: Sample inquiry tasks related to Ancient Rome.

ID	Question
1	Why did the first Romans settle on the hills?
2	Were the kings of Rome chosen by birth or by election?
3	What are the three social classes in ancient Rome?
4	Who were the Etruscan kings?

Protocol. As in [13], we defined search prompts to trigger interactions with SE in the classroom. Specifically, teachers asked their *primary five students to answer 12 questions related to a school subject* (i.e. Ancient Rome) using SE with (and without) enriched interfaces (sample prompts in Table 1). Teachers presented the task to children as a group, then each child completed the task individually. Study participants were exposed to all interfaces—they used each to answer 4 questions; interfaces were presented in random order. Upon completion of the inquiry task, we asked children to fill out a post-task questionnaire to elicit their preferences on mock-up interfaces: (i) **Favourite:** which interface is your favourite and why? (ii) **Helpful:** which interface would you use in the future to help you complete school assignments? and (iii) **Advised:** which interface would you suggest to your peers?. The study was approved by the local Ethics committee. As recent COVID-19 developments caused school closures, expert educators, who incorporated the proposed protocol as part of classroom instruction, conducted the study using online tools like Skype. Out of the 100 originally-recruited students only 31 completed the study as a result of the cognitive overload imposed on children by running the study completely online.

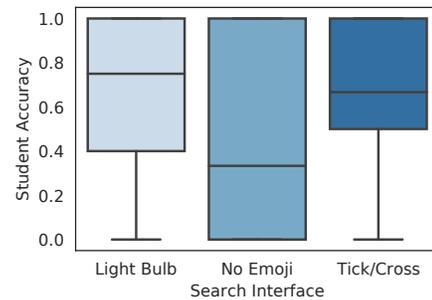
Data. For analysis purposes, we turn to responses to the post-task questionnaire for examining children’s preferences and the reasons behind them, in addition to their perceptions on the usefulness of emojis. We also rely on query logs generated as a result of children’s using the SE adapted to display the aforementioned interfaces (i.e. mock-ups) to complete the proposed search assignment. Query logs capture children’s explicit interactions with SE, and let us compute the values for electing the **Best Performing** interface.

Lenses. We rely on FHAB lenses to direct our exploration of collected data. Recall that we detailed each lens in Section 1.

3 RESULTS AND DISCUSSION

We discuss our analysis of post-task questionnaire responses and query long interactions, guided by FHAB lenses. Recall that we only considered the 31 of the participants (17 boys and 14 girls), who answered the 12 search prompts and the post-task questionnaire.

RQ1: Enriched Interfaces & Performance. We first examine whether interfaces that incorporate emojis offer children useful and usable cues of results’ (non-)relevance. Children perform the best on enriched interfaces (ANOVA one-sided test, $p < 0.001$). This is depicted in Figure 2; the median accuracy¹ of students in the low thirty percentile using a traditional interface, as opposed to median accuracy ranging from upper sixty to mid seventy percentiles on emoji-enriched interfaces. We also observe substantially different average accuracy per student, namely, Light Bulb: 0.63, No Emoji: 0.43, and Tick/Cross: 0.70. It is also remarkable that the box for the No Emoji interface exhibits a wide range of accuracy values, as opposed to the others. This serves as a baseline in our study,

**Figure 2: Box plot of student accuracy per interface.**

suggesting that depending on their prior knowledge of the subject or its difficulty students behave differently. However, we see that the boxes of the other two interfaces are considerably pushed up, indicating that the assistance and cues that the children get from the emojis actually help them distinguish relevant results, thus becoming more accurate in their clicks. This trend is appreciated in Figure 3, where the use of interfaces enriched either with Light Bulb or Tick/Cross emojis, as opposed to No Emoji interface, lead more students to perform more effectively. These results serve as strong indicators of the benefits of enriching interfaces with icons that can signal potentially relevant resources listed on SERP in response to children’s classroom-related inquiries. Our results pertaining to the classroom context further support those reported in [10], shining a light on the need to enrich interfaces with pictures to help children in their quest for online resources.

RQ2: Actions vs. Words. As illustrated in Figure 3, from our examination of responses to post-task questionnaire and query logs, it emerges that children mostly favour the interface enriched with Light Bulb emojis. Surprisingly, the majority selects the interface displaying the Tick/Cross emoji as the one they would use in the future to complete school search-related assignments. This is also the interface that children would most likely advise their peers to use for school-related information discovery assignments. This discrepancy aligns with reports in [1], in the sense that children mention favouring a particular interface, but when thinking of their use in a classroom setting, factors beyond preference are at play. In our study, “funny” and “new idea” are the most prevalent terms justifying why children prefer Light Bulb, whereas “clear” and “easy to understand” explain their choice for Tick/Cross.

For more thorough scrutiny, we turn to Figure 4, where we depict a pairwise analysis of FHAB lenses. Results reveal a 45% agreement between the B and F lenses; indicating that the interface that children explicitly state liking is not necessarily the one that prompts them to be more accurate, i.e. select more relevant results while reducing the number of (non-relevant and neutral) resources explored in SERP. Similarly, there is only a 42% agreement among B and H lenses, i.e. the interface that fostered best performance among children was not the one that they claimed was the most helpful. Further, 17 children favoured the Light Bulb interface, but only 11 would advise their peers to use it (A vs. F); 15 children identified Tick/Cross as the most helpful interface, but only two-thirds also thought it was their favourite (F vs. H); close to 30% of child participants who identified Light Bulb as the one they would suggest their peers to use would not use it themselves as they did not perceived Light Bulb as the most helpful interface (H vs. A). These results emphasise the fact that, when children are the main stakeholders, actions do speak louder than words.

¹Accuracy computed as the proportion of the number of clicks on resources known to be relevant (labelled relevant by expert educators) over the total number of clicks.

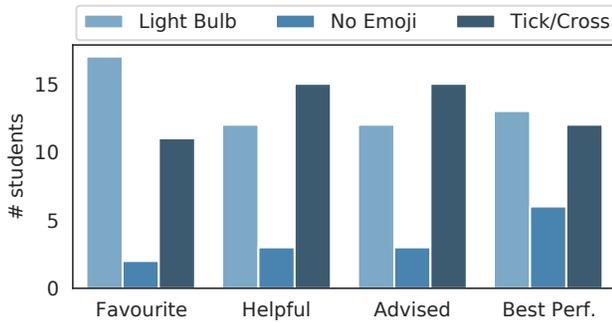


Figure 3: FHAB lenses on post-task questionnaires and query logs.

Do actions speak louder than words? Results' analysis reveal that just because children like an interface, it does not mean they would use it in the future or advise their classmates to use it. From insights across FHAB lenses (Figure 4), it is noticeable how asking to advise their peers brings children back to the formal school context and their assumption that in that setting, performance is more important than any aesthetic quality like fun or originality. This in turn would suggest that the FHAB lenses could serve as a tool to drive analysis of data gathered both in formal and informal contexts, and as such would allow exploration of how to design tools to support children searching for leisure. Moreover, our findings promote the use of emoji-enriched interfaces as conducive of accurate and effective search sessions; equally we highlight the importance of the aesthetic dimension in the design of interfaces for SE to be used by children, and specifically the inclusion of emotional elements to add fun and provide engagement during the search activity. While our participants seemed aware of and cared for the fun elements, their actions revealed how important performance was for them and how the formal context they were running their searches in influenced their choice of what interface to use in the future as well in what to recommend to peers. This information could only be elicited by combining *actions* with *words* (answers to post-task questionnaire) and inspect them through the four lenses.

Using post-task questionnaires to elicit preferences and behaviour from adult searchers is known to be problematic. Often users have confused ideas about the reasons behind a successful vs. a failed search [3]. This lack of clear understanding is more obvious and common when involving children [7, 13]. This inspired our use of the FHAB lenses to look into children's search behaviour more in-depth and understand what causes them to perceive the interface as usable and attractive. We posit the same lenses could be useful when running similar explorations with adults.

4 CONCLUSIONS AND FUTURE WORK

We used four lenses grounded on UX factors to analyse post-task questionnaires and query logs generated as a result of the study we conducted with 31 children in primary five. It is remarkable to see how poorly children fare when using a classic SERP, compared to when using an emoji-enriched one. Indeed, enriched interfaces better support completion of classroom search tasks, but which emoji-enriched interface to use remains uncertain. In other words, more research in this area is required, as differences in performance using interfaces augmented with Light Bulb vs. Tick/Cross were

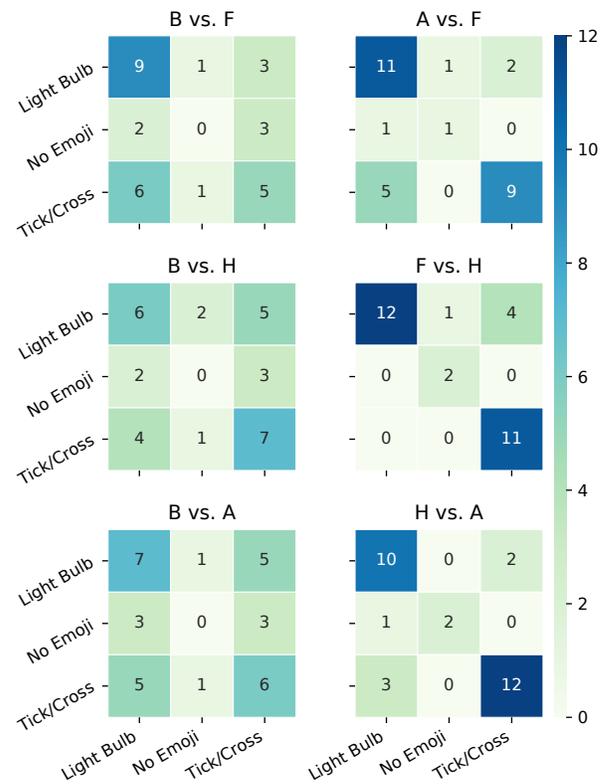


Figure 4: Pair-wise child agreement among FHAB lenses. Headers on sub-graphs indicate Y-axis lens vs. X-axis lens.

not statistically significant. An in-depth study of children's rationale to name a particular interface as their favourite is an immediate next step, as it seems that study participants prioritised an emoji being appealing over its ability to ease their information seeking activity, i.e. make the whole system perform better. Findings from our exploration also open future research paths related to investigating how good children are at determining which search tool is useful for them and, more importantly, what do we need to do to increase children's ability to judge for themselves what is relevant/useful to accomplish their goal. Certainly, this is of at most importance if we take into account that children receive little or no guidance or instruction when it comes to online searching.

As this was an initial exploration, we acknowledge some limitations: we compared the classic search interface without any relevance indicators against two mock-ups with explicit indicators; yet we do so, as children are familiar with plain interfaces on their preferred SE; running the study online was unavoidable but increased the cognitive overload required from children; lastly, we only used logs as proxies for explicit interactions.

Research outcomes we reported show how the four lenses proposed for our exploration allowed us to paint a rich picture characterising key aspects of children search behaviours in the classroom. Hence, the lenses can offer a template of sorts for future explorations, as they enable comparing and contrasting results across different proposed search tools and searchers. We believe this to be a valid contribution to the research community when looking for frameworks to analyse user data also considering adult searchers.

Acknowledgement. We want to thank all the children who took part of our study. We also appreciate the teachers who virtually administered our proposed protocol due to the pandemic.

REFERENCES

- [1] Mohammad Aliannejadi, Monica Landoni, Theo Huibers, Emiliana Murgia, and Maria Soledad Pera. 2020. Say it with Emojis: Co-designing Relevance Cues for Searching in the Classroom. In *Proceedings of the Joint Conference of the Information Retrieval Communities in Europe, CIRCLE*.
- [2] API | Microsoft Azure. Accessed March 19, 2020. *Bing Web Search API: Microsoft Azure*. <https://azure.microsoft.com/en-us/services/cognitive-services/bing-web-search-api>
- [3] Maram Barifah and Monica Landoni. 2020. Emotions associated with failed searches in a digital library. In *Proceedings of the Information Behaviour Conference, ISIC*.
- [4] Dania Bilal and Joe Kirby. 2002. Differences and similarities in information seeking: children and adults as Web users. *Information processing & management* 38, 5 (2002), 649–670.
- [5] Pia Borlund and Jesper W Schneider. 2010. Reconsideration of the simulated work task situation: a context instrument for evaluation of information retrieval interaction. In *Proceedings of the third symposium on Information interaction in context, IiX*. 155–164.
- [6] Judith H Danovitch. 2019. Growing up with Google: How children’s understanding and use of internet-based devices relates to cognitive development. *Human Behavior and Emerging Technologies* 1, 2 (2019), 81–90.
- [7] Allison Druin, Elizabeth Foss, Leshell Hatley, Evan Golub, Mona Leigh Guha, Jerry Alan Fails, and Hilary Hutchinson. 2009. How children search the internet with keyword interfaces. In *Proceedings of the International Conference on Interaction Design and Children, IDC*. 89–96.
- [8] Tatiana Gossen. 2016. *Search engines for children: search user interfaces and information-seeking behaviour*. Springer.
- [9] Infotopia. 2020. Infotopia: About Us. <https://www.infotopia.info/testimonials.html>. (Accessed October 13, 2020).
- [10] Hanna Jochmann-Mannak, Theo Huibers, Leo Lentz, and Ted Sanders. 2010. Children searching information on the Internet: Performance on children’s interfaces compared to Google. In *Proceedings of the International ACM SIGIR conference on research and development in Information Retrieval, SIGIR*. 27–35.
- [11] Ioannis Karatassis. 2017. WebSAIL: Computer-based Methods for Enhancing Web Search Literacy. In *Proceedings of the Conference on Conference Human Information Interaction and Retrieval, CHIIR*. 403–405.
- [12] Monica Landoni, Theo Huibers, Emiliana Murgia, Mohammad Aliannejadi, and Maria Soledad Pera. 2021. Somewhere over the Rainbow: Exploring the Sense for Relevance in Children. In *Proceedings of the European Conference on Cognitive Ergonomics (ECCE)*.
- [13] Monica Landoni, Davide Matteri, Emiliana Murgia, Theo Huibers, and Maria Soledad Pera. 2019. Sonny, Cerca! evaluating the impact of using a vocal assistant to search at school. In *Proceedings of the International Conference of the Cross-Language Evaluation Forum for European Languages, CLEF*. 101–113.
- [14] Mengyang Liu, Yiqun Liu, Jiaxin Mao, Cheng Luo, Min Zhang, and Shaoping Ma. 2018. "Satisfaction with Failure" or "Unsatisfied Success" Investigating the Relationship between Search Success and User Satisfaction. In *Proceedings of the World Wide Web Conference, WWW*. 1533–1542.
- [15] Ion Madrazo Azpiazu, Nevena Dragovic, Oghenemaro Anuyah, and Maria Soledad Pera. 2018. Looking for the Movie Seven or Sven from the Movie Frozen? A Multi-perspective Strategy for Recommending Queries for Children. In *Proceedings of the Conference on Human Information Interaction & Retrieval, CHIIR*. 92–101.
- [16] Ivonne Moncarca, Franceli L Cibrian, Angel Mendoza, Gillian Hayes, and Monica Tentori. 2020. Why doesn’t the conversational agent understand me? a language analysis of children speech. In *Adjunct Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing and International Symposium on Wearable Computers, UbiComp/ISWC*. 90–93.
- [17] WizeNoze. 2020. WizeNoze: WebforClassrooms. <https://webforclassrooms.com/>. (Accessed October 20, 2020).
- [18] Svetlana Yarosh, Stryker Thompson, Kathleen Watson, Alice Chase, Ashwin Senthilkumar, Ye Yuan, and AJ Bernheim Brush. 2018. Children asking questions: speech interface reformulations and personification preferences. In *Proceedings of the ACM Conference on Interaction Design and Children, IDC*. 300–312.
- [19] Xiaojun Yuan and Ning Sa. 2020. Speak to Me: Interacting with a Spoken Language Interface. In *Proceedings of the International Conference on Human-Computer Interaction, HCI*. Springer, 194–211.
- [20] Jun Zhao, Ge Wang, Carys Dally, Petr Slovak, Julian Edbrooke-Childs, Max Van Kleek, and Nigel Shadbolt. 2019. I make up a silly name’ Understanding Children’s Perception of Privacy Risks Online. In *Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI*. 1–13.