

How Kids See Search: A Visual Analysis of Internet Search Engines

Nicholas Vanderschantz
Department of Computer Science
The University of Waikato
New Zealand
vtwoz@waikato.ac.nz

Annika Hinze
Department of Computer Science
The University of Waikato
New Zealand
hinze@waikato.ac.nz

Through previous studies into children's internet search practice, we have gained insight into the taught strategies, information behaviour, and common errors children experience while searching. This paper analyses the visual structure of commonly-used internet search engines (ISEs) to explore how the interface and interaction design of ISEs may influence the search practices of children. Common features of ISEs are identified and the effects of query construction techniques on the visual presentation of information are reported. We use our observations to provide guidelines for the design and development of ISEs for children.

web search, children, digital information behaviour, internet search engines, information search

1. INTRODUCTION

Children use internet search engines (ISEs) as their primary information source when searching for information at home and at school (Vanderschantz et al. 2014b). In this paper, we assess how the design of contemporary ISEs may influence the search practices of children. We firstly aim to identify and name common features of search engine results pages (SERPs) of ISEs. We secondly show the effects of typical children's query construction strategies on the visual presentation of information in a search engine results page. The findings of our study will assist in providing visual guidelines for the design of ISEs for children aged 9 to 13 years old.

2. RELATED WORK

Studies into children's information search focussed on printed artefacts (e.g. Moore 1995) or older ISEs (e.g. Bilal 2000). Recent investigations included analysis of information search logs (e.g. Duarte Torres et al. 2010) and qualitative and quantitative analysis of children's search habits at home (e.g. Druin et al. 2010), and in educational settings (Vanderschantz et al. 2014b,a). A widely accepted Information Retrieval (IR) problem for children is their lack of vocabulary or domain knowledge, which likely creates spelling issues and impedes understanding of the correct use of search interfaces (e.g. van der Sluis and van Dijk 2010; Bilal 2000). Children also have potentially-flawed mental models

of ISEs (Norman 1983). They use ISEs that are predominantly designed for adults and which do not suit their information-seeking needs. A common response to the issue are specialised child-centred IR systems (e.g., Druin et al. 2003; Gossen et al. 2012; Lingnau et al. 2010). These are often research-based prototypes without ongoing support, while many commercial child-specific ISEs have since disappeared from the market. We found that no dedicated child-centred systems are used in NZ classrooms (Vanderschantz et al. 2014b,a).

Our previous studies with children and teachers (Vanderschantz et al. 2014b; Vanderschantz 2016; Vanderschantz and Hinze 2017b) explored typical inquiry tasks in NZ classrooms. Beyond using keywords, children are taught to use *Query Qualifiers* (e.g., "facts" or "for kids") and *Query Refiners* (e.g., further topic words or natural language). In our studies and the literature (e.g. van der Sluis and van Dijk 2010; Bilal 2000), children are reported to use questions with and without question marks, and full sentences with conjunctions.

3. STUDY METHOD

We conducted a lab-based study in which a series of child-typical search queries was executed. We selected search engines that would be encountered by children in an educational setting: Google, Bing, and Yahoo!, and Google with Reading Level Filter set

to 'Basic'. For simplicity, we refer to this setting as a separate, fourth search engine (Google-B).¹ We used the Chrome web browser in Incognito mode with no user style-sheets activated. The viewable area of the browser was 1208 pixels by 1048 pixels. The horizontal position of 1208 pixels from the top of the browser window will be referred to as the *fold*. Unless explicitly stated, we did not scroll nor did we record screenshots of the information below the fold.

Query Construction. We aimed to replicate the process of a typical inquiry task in NZ classrooms by choosing a common topic of investigation ("Mount Everest"). To study the factors that influence the presentation of search results and the impact this will have on children's ability to use SERPs, we tested 43 variants of a search query about 'Mount Everest'. When developing appropriate test queries, we applied both query qualifiers and query refiners, as well as capitalisation, abbreviation, punctuation and search operators. The queries were thus constructed in a way that replicates typical search strategies of children in Years 5 to 8 of NZ schools (aged 9 to 13 years). Selected example queries are Mount Everest, Mount Everest facts, mount everest for kids, mount everest height, and What is the height of mount everest.²

Evaluation Criteria. We evaluated the potential impact of the information presentation and the visual features of ISEs on children's search. In particular, we identified graphic and typographic features and we now discuss their potential impact on children's ability to read and find information when triaging and searching. We noted differences in search result list orderings, but did not speculate on possible implications of list ordering nor the inherent quality of the returned websites. Our intention was not to reverse engineer the algorithms used by these search engines but to provide insights into the presentation of information to children.

4. RESULTS

Here we report the findings of our visual analysis that resulted from our query constructions tests.

4.1. Conventions and Definitions

We use the layout references labelled (A) to (J) in Figure 1 and Figure 2. We identify two categories of

¹Google and the Google logo are registered trademarks of Google Inc., used with permission. Bing and the Bing logo are registered trademarks of Microsoft Inc., used with permission. Yahoo! and the Yahoo! logo are registered trademarks of Yahoo! Inc., used with permission. This declaration holds for all of the content of this paper including text and images used henceforth.

²For a full list of queries and further details see Vanderschantz and Hinze (2017a)

visual presentation elements in search results pages (SERPs). *Entry-level units* are the typical search results entries (e.g., (A)). These were found to be very similar across all four ISEs, consisting of title, URL, and descriptor (see (D), (E), (F)). *Block-level units* are visually separated from other information using techniques such as borders and background colours; we identified pull-boxes (e.g., (B), (C)) and sidebars (e.g., (G)). Note that we refer only to entries and visual presentations above the fold (see (H)).

4.2. Search Results Page Presentation

All four search engines used similar overall SERP layout and we observed strong visual similarities. All ISEs used a left-aligned page layout with top left branding next to the search box (see Figure 1). All four required scrolling to review the entire SERP list. Google, Google-B and Bing presented a two-column layout, while Yahoo! used a three-column layout. The two-column layout appeared more visually open and suffered less from crowding. Open and uncrowded SERPs will assist with effective eye paths that will be more beneficial for children who easily become distracted by complex information presentation.

Text Presentation. Google, Google-B and Bing used a white background colour for both the page and sidebars, while Yahoo! used a light grey background colour. The text and background colour contrasts are likely appropriate for readability and legibility for children. All four ISEs use a similar number of characters per line, a single line of text for titles, and two lines for descriptors. Keywords are displayed in bold in search result titles, URLs, and descriptions. Entry titles appear in blue (see (D) in Figure 1), URLs appear in green ((E) in Figure 1), and descriptors in grey (see (F)). All ISEs use the same type size for URL and descriptor text.

Text spacing is important for leading the eye around a document in an effective and efficient manner. On-screen reading by children is affected by the text size, line-length, interlinear space (i.e., horizontal space between lines) and paragraph space (i.e., space between blocks of text) as well as margins and gutters (i.e., space outside and between blocks of text). Yahoo! provided the most generous space between SERP list entries, followed by Google and Google-B and then Bing. Google, Google-B and Bing had greater interlinear space between title and URL than Yahoo!. The spacing between lines of text will help with the readability of individual lines and will improve the skimmability and scannability of SERP entries. Both Bing and Yahoo! use underlined text for the titles of each entry, which decreases the interlinear spacing and therefore scanning for the titles of entries may be less efficient for children.

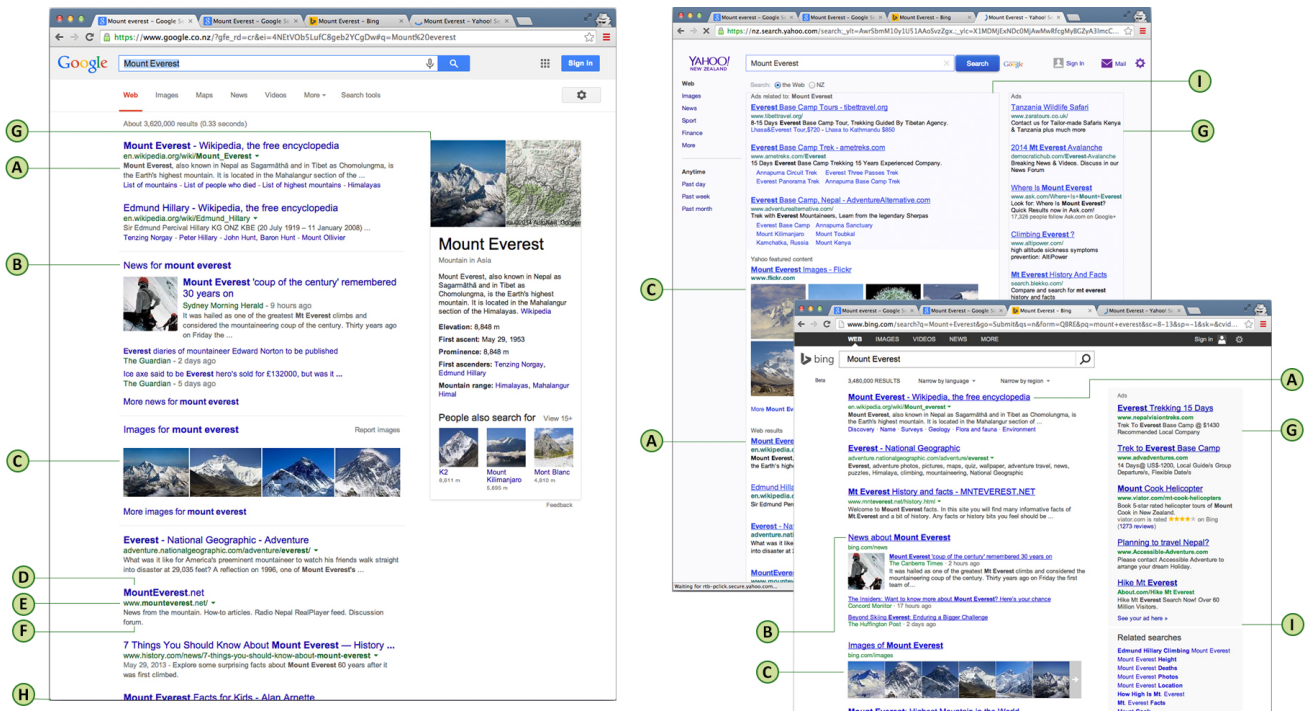


Figure 1: Google (L), Yahoo! (Top R), Bing (Bottom R). Google and the Google logo are registered trademarks of Google Inc., used with permission. Yahoo! and the Yahoo! logo are registered trademarks of Yahoo! Inc., used with permission. Bing and the Bing logo are registered trademarks of Microsoft Inc., used with permission.

Ad-blocks. When discussing advertising on web pages, one teacher in our interviews in (Vanderschantz 2016) explicitly reported: websites cluttered with advertisements hinder [children] when finding information. Inclusion of advertising links, is therefore likely detrimental to successful information search by children. For the queries performed, Bing and Yahoo! both showed ad-blocks in Position 1 of the search result list (see ① in Figure 1). Yahoo! returned three advertisements in Position 1 for every search conducted for our study, while Bing only sometimes displayed advertisements. Bing used a light-green background-colour with a grey right-hand border with the title ads in the top right corner. Yahoo! used a light-grey background-colour and the title advertisements related to: Bing and Yahoo! both also used the sidebar to show advertisements.

To force advertisements to appear in Google we ran the search query travel mount everest, (not one of the initial 43 queries). Using this query we were able to return two advertisements in both Google and Google-B. Advertisements in Google appeared in pull-boxes in Position 1. Each sponsored link was marked with a small yellow graphic next to the URL.

Image-blocks and Video-blocks. Image-blocks were included in SERPs by all four ISEs. Google (including Google-B) and Bing presented 4 and 6 images, respectively, in a carousel strip, while Yahoo! presented 8 images in a grid. Google encapsulated

image-blocks (see ③ in Figure 1) with a top and bottom border to visually demarcate these from the run of typical entries. Bing added additional space before and after an image-block to visually separate this from typical entries. Pull-boxes used for videos were only noted in Bing and Google (and Google-B). Bing clearly marked a video with a pull-box containing multiple videos in a strip similar to how Bing displays image-blocks. Google, however, simply included single videos, with a video *still-frame* as a clickable icon to the left of an entry.

Info-blocks and Sidebars. Google and Google-B were the only ISEs to incorporate a unique pull-box which we have called info-blocks (see ④ in Figure 2). These were noted only when a specific question or recognizable fact was used as the search query. Info-blocks returned the answer to the question within the SERP and supplied a link to the source website as well as images where relevant. This search-related-information always included an image and a short list of factual details. The provenance of some of the information is clear and can be tracked by way of the link text provided, while some of the information's provenance is more difficult to identify. We noted that Google and Google-B also used sidebars to list information related to a search. Sidebars we identified as the right-most column of a SERP. Yahoo! and Bing both included advertisements in the top-most portion of the sidebar and never info-block content akin to that

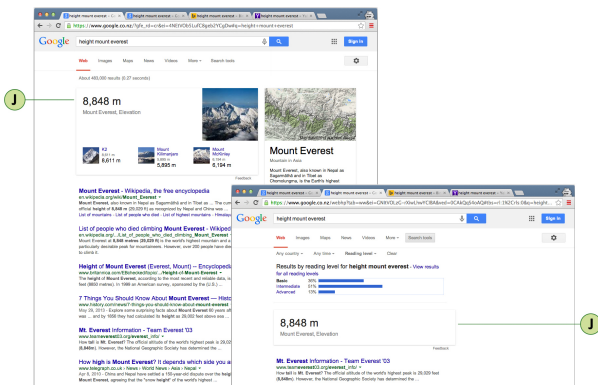


Figure 2: Info-boxes Google (L) & Google-B (R). Google & the Google logo are registered trademarks of Google Inc.

of Google. Yahoo! typically displayed six to eight advertisements while Bing often displayed three advertisements. Bing included a lower section to its sidebar containing related-searches above the fold that will assist users with query reformulation.

News-blocks. Google and Bing both included news-blocks (see (B) in Figure 1). These did not appear in Google-B or Yahoo! for any of our searches. We acknowledge that news-blocks serve a greater informational purpose than ad-blocks, for example. Yet, these may still obscure informational SERP links and may be detrimental to successful information search by children. Specific investigation of the impact of all of these block level items is warranted.

Related Searches. Related searches features of ISEs were investigated because in previous studies the children (Vanderschantz et al. 2014b) and teachers (Vanderschantz 2016) had discussed difficulties the children encounter when reformulating searches. To review the use of Related Searches for all ISEs we did need to scroll the browser to review information that was below the fold. That related searches are demoted to below the fold and some distance from the search box may mean children do not associate this tool with query reformulation. Google, Yahoo!, and Bing placed related-searches at the bottom of the search result list (below the fold). Bing was the only ISE to also clearly display related-searches in the sidebar (above the fold). Google-B did not display related-searches for any of our search queries. Related searches were present for all searches in Bing and Google, but not for Yahoo!. When related-searches are not present in a Yahoo! SERP, advertisements are placed in the same space, which may cause confusion.

4.3. Query Construction Effects

We observed that minor changes in query construction can have significant effects on both result list as well as SERP presentation.

Keyword (KW) vs. Natural Language (NL) Queries. All four ISEs showed differences in very early entries in the SERP list (often Position 1 or 2) when comparing a keyword search to natural language search for either sentence or question constructions. In Bing, news-blocks were only displayed for KW searches, not NL searches. In Yahoo! video-blocks were only used for KW searches. When Natural Language Sentences (NLS) and Natural Language Questions (NLQ) were used, Bing and Yahoo! both presented more advertisements. Google and Google-B resulted in the fewest differences when NLQ are compared to keyword queries and do not produce advertisements for these queries.

Natural Language Sentences (NLS) vs. Natural Language Questions (NLQ). Using NLS vs NLQ resulted in differences occurring at Position 6 or 7 across all ISEs. NLS triggered news-blocks and video-blocks in Bing.

Punctuation and Capitalisation. seemed to have little or no impact on visual presentation or search results. The use of these query construction techniques should not hinder children's ability to find websites to visit from the returned SERP lists.

Phrase Ordering. This refers to the permutation of keywords within the query (e.g., mount everest facts vs. facts mount everest). We observed an effect on the SERP lists of all four ISEs, typically at early positions such as Entry 2 or 3.

Abbreviations. Different to the full query (e.g., mount everest), Google and Google-B did not show any news items for the queries using abbreviations (e.g. mt everest). Bing removed news-blocks and video-blocks when an abbreviation was used resulting in more full entries above the fold compared to non-abbreviations. Yahoo! produced only differences in the content of the advertising.

Search Operators. Although Boolean search operators, such as and, + and -, were not specifically taught to children in great detail, teachers use these in model searches. Google and Google-B were not affected by the use of query operators. Bing and Yahoo! displayed more advertising when an operator was used. The content of the sidebars in Bing was different when using query operators.

Query Qualifiers. The children in the schools that we studied are taught to use search qualifiers such as facts and for kids. Adding these phrases to a query is intended to result in websites whose language or structure are appropriate for children. The use of a query qualifier produced noticeable effects in the entry list with changes to the first

entry for Google, Google-B and Bing. The addition of qualifiers resulted in both Google and Bing removing news-blocks from their SERP lists. Google further dropped sidebars. Bing and Yahoo! dropped image-blocks and Bing also dropped video-blocks. The addition of query qualifiers resulted in more 'Basic' search results in Google-B.

5. DISCUSSION

There are no closely-related studies with which to compare to our results. However, our research relates to the area of *Visual Aesthetics* (Tractinsky 2013). De Angeli et al. (2006) provide empirical evidence that aesthetically-considered design of interactive technology can increase users pleasure and engagement, furthermore, Moshagen stated that visual aesthetics may improve performance and thereby compensate for usability problems (Moshagen et al. 2009, p. 1317). The development of this research area as motivation for further investigations of children's information seeking interfaces. We identified a number of visual conventions used by these search engines yet little empirical evidence is presently available regarding the effects of these visual features for web searchers and no evidence is reported in the literature regarding children's use of these visual features of ISEs. Future studies on HCI, and IR that assess the effectiveness of these visual features for children and adults are required.

In previous studies (see Vanderschantz et al. 2014b; Vanderschantz 2016), both children and teachers reported on skimming and scanning behaviour being a necessary feature of children's use of SERP. While eye movement studies to evaluate commercial websites are now commonplace, we found no empirical studies based on SERPs. Eye movement studies (e.g. Buscher et al. 2009; Hervet et al. 2011) of web pages (differing from web search engines) often show the eye is drawn to visual stimuli such as images and logos. Thus, future investigation with children could gauge the effect of SERP features such as pull-boxes. We observed that the use of pull-boxes resulted in a decreased number of search result entries visible above the fold. The number of entries above the fold may be a contributor to ease of skimming, scanning and information triage by children, which may require further investigation.

Bilal (2013) has begun investigating the effectiveness of the Google-B and in our previous studies (see Vanderschantz et al. 2014b; Vanderschantz 2016), children and teachers reported the use of Google-B. At the time of writing, this reading level filter feature has been removed from Google and no replacement has been incorporated. Should a similar

feature be integrated into Google or another ISE the use of that feature by children requires investigation.

The rate of technology advancement must be considered during the reporting of such a study, while acknowledging the research insight that is independent of commercial considerations. When repeating a selection of the searches at the time of writing, fewer advertisements were noted and less space is used for advertisement blocks by Bing and Yahoo!. All search engines still display images and videos in pull-boxes, yet in slightly different visual styles. Google is still the only ISE to incorporate pull-boxes for information. These modifications by all manufacturers is evidence of the rapid change and supports the continual and progressive research into what is required to assist children and adults in searching for information. We believe these design changes by the ISEs do not compromise our findings regarding what is necessary in an ISE for children and how query structure effects SERP presentation.

6. SUMMARY & CONCLUSION

This paper reported selected results of a lab-based study exploring the visual presentation of internet search results and implications for children's search. While visual design and presentation of SERP result lists was found to be fairly consistent across search engines, we identified visual differences in SERP design and list ordering created by alternative query structures and phrasing of similar queries. We draw the following conclusions from our study:

1. Some query construction techniques may assist with returning result pages that are visually advantageous for children. For example,

- Using abbreviations resulted in more entries above the fold due to fewer news-blocks and less advertising. Abbreviations also appeared produce a larger number of 'Basic' results in Google-B compared to the full spelling.
- Use of query qualifiers such as *facts* and *for kids* seemed to have overall positive effects on keyword searches by reducing advertising and news-blocks (more results above the fold).

2. Design could better support children's information search through the following design guidelines.

- Explicit support for query reformulation would benefit children. Emphasis of alternative query constructions in the form of related searches could be positioned high on SERP to encourage use of this query reformulation tool.
- Non-result-list content should be visually different from results list links to assist skimming

and scanning. Pull-boxes and sidebars that contain information related to the search query can serve the needs of young searchers by providing useful resources to complement the websites presented in the search lists.

- Clear visual differentiation of advertising entries is required to assist young readers in identifying sponsored links compared to search results. Advertising was used less with natural language queries suggesting these might prove a successful search technique when looking to avoid advertisements.
- Designers are recommended to visually differentiate sidebars, pull-boxes, and advertising through use of borders, background colour differences, typographic differences and increased white space around these features.

Our work reported here finally identifies the need for further investigations into how the identified interface elements affect children's information search and how the presentation differences created by alternative query formulations impacts children's ability to find information.

REFERENCES

- Bilal, D. (2000). Children's use of the yahoooligans! web search engine: I. cognitive, physical, and affective behaviors on fact-based search tasks. *Journal of the Association for Information Science and Technology*, 51(7):646–665.
- Bilal, D. (2013). Comparing google's readability of search results to the flesch readability formulae: A preliminary analysis on children's search queries. *American Society for Information Science and Technology*, 50(1):1–9.
- Buscher, G., Cutrell, E., and Morris, M. R. (2009). What do you see when you're surfing?: using eye tracking to predict salient regions of web pages. In *SIGCHI conference on human factors in computing systems*, pages 21–30. ACM.
- De Angeli, A., Sutcliffe, A., and Hartmann, J. (2006). Interaction, usability and aesthetics: what influences users' preferences? In *Designing Interactive systems*, pages 271–280. ACM.
- Druin, A., Bederson, B. B., Weeks, et al. (2003). The international children's digital library. Technical Report CS-TR-4433 UMIACS, University of Maryland.
- Druin, A., Foss, E., Hutchinson, H., Golub, E., and Hatley, L. (2010). Children's roles using keyword search interfaces at home. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 413–422. ACM.
- Duarte Torres, S., Hiemstra, D., and Serdyukov, P. (2010). An analysis of queries intended to search information for children. In *Symposium on information interaction in context*, pages 235–244. ACM.
- Gossen, T., Nitsche, M., and Nürnberg, A. (2012). Knowledge journey: A web search interface for young users. In *Symposium on Human-Computer Interaction and Information Retrieval*, page 1. ACM.
- Hervet, G., Guérard, K., Tremblay, S., and Chtourou, M. S. (2011). Is banner blindness genuine? eye tracking internet text advertising. *Applied cognitive psychology*, 25(5):708–716.
- Lingnau, A., Ruthven, I., Landoni, M., and Van Der Sluis, F. (2010). Interactive search interfaces for young children—the puppyir approach. In *Int. Conference on Advanced Learning Technologies (ICALT)*, pages 389–390. IEEE.
- Moore, P. (1995). Information problem solving: A wider view of library skills. *Contemporary educational psychology*, 20(1):1–31.
- Moshagen, M., Musch, J., and Göritz, A. S. (2009). A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance. *Ergonomics*, 52(10):1311–1320.
- Norman, D. A. (1983). Some observations on mental models. *Mental models*, 7(112):7–14.
- Tractinsky, N. (2013). Visual aesthetics. In *The encyclopedia of Human Computer Interaction*. Interaction Design Foundation.
- van der Sluis, F. and van Dijk, E. M. A. G. (2010). A closer look at children's information retrieval usage: Towards child-centered relevance. In *Workshop on Accessible Search Systems, at ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2010)*, pages 3–10.
- Vanderschantz, N. (2016). *KidsQuestions: Assisting children's digital information seeking*. PhD thesis, Computer Science, University of Waikato, New Zealand.
- Vanderschantz, N. and Hinze, A. (2017a). Do Internet Search Engines Support Children's Search Query Construction: A Visual Analysis. Technical Report 02/2017, University of Waikato, Department of Computer Science.
- Vanderschantz, N. and Hinze, A. (2017b). A study of children's search query formulation habits. In *2015 British HCI Conference*. ACM.
- Vanderschantz, N., Hinze, A., and Cunningham, S. J. (2014a). Current educational technology use for digital information acquisition by young new zealand children. In *Australasian Computer Science Conference*, pages 125–134. ACS.
- Vanderschantz, N., Hinze, A., and Cunningham, S. J. (2014b). "Sometimes the internet reads the question wrong": Children's search strategies & difficulties. *American Society for Information Science and Technology*, 51(1):1–10.