Understanding the Needs of Searchers with Dyslexia

Meredith Ringel Morris[1], Adam Fourney[1], Abdullah Ali[1,2], Laura Vonessen[1,2]
[1]Microsoft Research, [2]University of Washington
{merrie, adamfo}@microsoft.com, xyleques@uw.edu, laurav4@cs.washington.edu

ABSTRACT
As many as 20% of English speakers have dyslexia, a language disability that impacts reading and spelling. Web search is an important modern literacy skill, yet the accessibility of this language-centric endeavor to people with dyslexia is largely unexplored. We interviewed ten adults with dyslexia and conducted an online survey with 81 dyslexic and 80 non-dyslexic adults, in which participants described challenges they face in various stages of web search (query formulation, search result triage, and information extraction). We also report the findings of an online study in which 174 adults with dyslexia and 172 without dyslexia rated the readability and relevance of sets of search query results. Our findings demonstrate differences in behaviors and preferences between dyslexic and non-dyslexic searchers, and indicate that factoring readability into search engine rankings and/or interfaces may benefit both dyslexic and non-dyslexic users.

Author Keywords
Dyslexia; reading disabilities; learning disabilities; Web search; search engine.

ACM Classification Keywords
H.3.3. Information Search and Retrieval; K.4.2. Social issues: Assistive technologies for persons with disabilities.

INTRODUCTION
Dyslexia is a language disability that is estimated to impact 15 – 20% of English speakers [17] (incidence rates vary by language [16]). People with dyslexia may experience a range of challenges relating to reading and spelling, such as slow reading rate, low reading comprehension, difficulty reading long passages, difficulty organizing language, difficulty remembering written or numerical information, and spelling challenges [17]. Because dyslexia is a spectrum disorder, not all people with dyslexia experience the same subset of challenges, and the severity of these challenges may vary, as well. Dyslexia is currently diagnosed through a comprehensive professional evaluation of abilities such as decoding, spelling, phonological processing, automaticity, reading comprehension, and vocabulary knowledge [18], though researchers are working toward developing simpler diagnostic tools (e.g., [34]).

Given the high incidence of dyslexia, and the pervasiveness and importance of web search as a method of finding information in modern life, it is important to consider whether the process of using a search engine to find information on the web is sufficiently accessible to people with dyslexia, and how this accessibility may be improved.

To better understand this issue, we conducted interviews with ten adults with dyslexia, discussing their experiences conducting web searches, including challenges they encountered and work-arounds they employed. Based on the themes that emerged in these interviews, we then designed and deployed an online questionnaire to gather further information from 80 dyslexic and 81 non-dyslexic adults about their experiences with web search. Finally, we selected ten challenging search queries, as identified by the dyslexic respondents of our questionnaire, and deployed an online study in which 174 adults with dyslexia, and 172 adults without dyslexia rated the relevance and readability of the corresponding search results. Taken together, these data directly address our two primary research questions:

RQ1: What challenges do people with dyslexia encounter when searching for information online, and what mitigation strategies do they employ?

RQ2: To what extent do the behaviors and preferences of dyslexic searchers differ from non-dyslexic searchers?

Additionally, our work contributes both user interface and algorithmic design suggestions to improve search engine accessibility for people with dyslexia.

RELATED WORK
Although dyslexia and other reading disabilities are quite common [17], there is relatively little research investigating how general-purpose algorithms or user interfaces might be specialized for this audience, though that is beginning to change. For instance, Rello et al. proposed spellchecker features targeted at the kinds of errors people with dyslexia may make [36], and created an online game that can help diagnose dyslexia [34]. Microsoft recently announced Learning Tools for OneNote [onenote.com/learningtools], which offers features for users with reading challenges, such as alternative spacing and fonts, and the ability to hear text read aloud. There is also a proliferation of tablet and smartphone apps that aim to help people with dyslexia, particularly apps targeted toward educational games and activities to improve literacy skills for children with reading

CHI 2018, April 21–26, 2018, Montreal, QC, Canada © 2018
Copyright is held by the owner/author(s).
Publication rights licensed to ACM.
ACM 978-1-4503-5620-6/18/04…$15.00
https://doi.org/10.1145/3173574.3173609

disabilities (e.g., Lectio, Easy Spelling Aid, Ghotit Real Writer, Dyslexia Quest, etc.).

However, there is little exploration of the impact of dyslexia on Web use, beyond the accessibility of certain fonts [35], discussion of the need to consider dyslexia when updating Web accessibility guidelines [10, 37], or studying the impact of challenges associated with dyslexia (reading difficulties, working memory difficulties, organizational difficulties) on general web navigation behaviors [2].

To our knowledge, the only two past studies of search behavior by people with dyslexia are the work of Berget et al. [4, 5] and MacFarlane et al. [25, 26]. Berget found that the absence of spelling-assistance functionality in a specialized library catalogue search tool hindered dyslexic users [5], but observed no difference in the use of Google's auto-complete functionality in a lab study comparing dyslexic and non-dyslexic searchers (for Norwegian-language queries). This suggests that well-designed interface features (e.g., autocomplete to assist with spelling challenges) may mitigate some dyslexia-related performance issues [4]. MacFarlane et al. compared TREC task performance for a small set of dyslexic and non-dyslexic searchers, finding dyslexic users required more search iterations [25] and had more difficulty eliminating irrelevant results [26]. Additionally, eye-tracking studies have shown that dyslexic and non-dyslexic users have different search-result scanning patterns [3, 27].

Research on information retrieval for general audiences also informs our investigation. Some reading level features have been proposed in the past for use in search ranking, with a focus on search engines for children [8, 21, 23]. Likewise, Ivory et al. [19] examined how properties of web pages related to overall page quality. Using their Find It If You Can search game, Ageev et al. found that triage and information extraction (as opposed to query formulation) were the skills that separated successful from unsuccessful searches [1]. Our work builds on this prior work by examining how reading level and other web page properties may impact triage and information extraction by dyslexic searchers.

INTERVIEW STUDY

To better understand the current experience of search engine use for adults with dyslexia, and to answer our first research question, we conducted an interview study.

Interview Study: Participants

We recruited US.-based participants with Facebook ads that invited people to participate in a "dyslexia research study" and directed them to fill out a brief screener collecting information on age, location, native language, dyslexia diagnosis, and search engine use. We used this screener to identify participants who were adult, native English speakers with dyslexia living near our lab that used web search engines at least several times per week. We scheduled phone interviews with participants, and continued recruiting until we reached a point of diminishing returns, i.e. when we began hearing the same themes without learning significant new information, in this case, after ten interviews.

Our participants were 39 to 69 years old (mean 46.6); eight were female, two male. Participants had a wide range of educational and socio-economic backgrounds, and dyslexia impacted their lives to varying degrees; some had dropped out of high school due to learning challenges, while others had obtained graduate degrees. Occupations included teacher, writer, social media guru, stay-at-home parent, student, retiree, self-employed, and paralegal.

Interview Study: Method

Interviews were conducted by phone in January and February 2016, and lasted about 30 minutes. Participants received a $20 Amazon gift card as a gratuity. Two researchers jointly conducted all ten interviews, using a semi-structured interview technique in which a core set of questions were asked of all participants, but additional questions were included on-the-fly to explore interesting topics raised in participants' answers. The researchers recorded the phone calls and took detailed notes.

The core interview questions first collected demographic details: age, gender, occupation, education, age at dyslexia diagnosis, specific symptoms experienced, and search engine use (search frequency, preferred sites, preferred devices). Next, the interview employed a recent critical incident technique [11], asking participants to describe their most recent web search in detail (i.e., the search need, the process used, the outcome, any challenges encountered). We also asked participants what they found most challenging about online information-seeking, whether they think dyslexia specifically makes search more challenging for them (and if so, how), and to describe strategies they used to overcome dyslexia-related search challenges. We concluded by asking participants whether and how they used specific technologies that might be helpful to people with dyslexia, including voice search, image search, modified fonts, ad blockers, and auto-completion. We also asked if participants had any ideas regarding how search engine companies might make it easier for people with dyslexia to find information on the web.

Two researchers employed standard qualitative analysis techniques to analyze the interviews, using variations of open coding and affinity diagramming [22] to iteratively identify themes in the data until reaching a comprehensive and mutually-agreed-upon set of themes.

Interview Study: Findings

Our interviews revealed that many people with dyslexia experience significant challenges in finding information in all three stages of the search process: query formulation, search result triage, and information extraction.

Query Formulation

Difficulties with spelling and reading made query formulation particularly challenging for our participants. Many reported that their misspellings of query terms were so far off from the correct spelling that search engines' spelling

correction functionality did not help them (e.g., it would spell-correct to a word other than the target query). For instance, P6 described trying to find information about the spider species tarantula but she misspelled it as *trantla* which the search engine did not successfully correct. Both P6 and P7 described dropping syllables from words, which seemed to pose particular challenges for auto-correction. P2 reported using other spellcheck software (e.g., her word processer) to first spellcheck her queries before copying and pasting them into her browser (the word processer spellchecker was preferred because it offered more options for what the target word might be, rather than the search engine's autocorrecting to a single option). Word substitutions were another dyslexic symptom not well-supported by autocorrect functionality; for example, P1 recalled a time when she was trying to find information about the household appliance *garbage disposal*; however, one of her symptoms is that she mixes up similar words. She entered the query *garbage exposer,* and the search engine's speller was unable to correct the term.

Because of reading difficulties, participants reported that it was challenging for them to verify whether automatic spelling correction had helped or hindered their query. Half reported using image search results to help with query formulation or verification. Viewing the image search results was a way to verify that the query they issued (or the auto-correction suggested by the search engine) matched the concept they had in mind.

Voice input was highly valued by searchers with dyslexia as a means to circumvent spelling challenges (though verifying that a voice search had been interpreted correctly still entailed reading challenges). Eight of our participants reported an affinity for voice input when available; on mobile phones, all major platforms support voice queries (e.g., via agents like Siri or Cortana). However, on PCs and laptops, participants expressed displeasure that voice input search was not always available (or not available by default, so therefore undiscovered by participants). P4 was so dependent on voice input for spelling that on his PC he used the Dragon Naturally Speaking dictation software, and would dictate query terms into another document and then copy and paste them into his search engine.

Search Result Triage

When conducting informational [7] searches, our participants reported challenges in determining which websites listed in the SERP (Search Engine Results Page) they should click on. This finding mirrors that of MacFarlane et al., who reported that dyslexic searchers have difficulty eliminating non-relevant documents when doing search result triage [26]. However, our participants reported that they considered a range of factors related to accessibility and readability, in addition to relevance, when selecting search results. These factors included:

1. Multimedia: Participants reported preferring pages that contained multimedia (e.g., images, videos) in addition to text (to help support their understanding in the event of reading difficulties). For example, P7 reported using the image search results as a way to find webpages that contained explanatory images, and P8 also noted that the image search results helped her find websites that were more "friendly" to dyslexic users.

2. Reading Level: Searchers with dyslexia see value in being able to quickly identify pages whose reading level matched their abilities (e.g., P8 noted having trouble with "long words"; P4 reported having to use Dragon language software to read difficult words aloud to him).

3. Clutter: For this audience, it was particularly valuable to identify pages with few or no advertisements or other distracting content. Half of participants noted that extraneous content made it more difficult for them to read effectively. Websites that divide up a single article across multiple pages also added interpretation complexity; simplicity in structure and visual layout was preferred.

4. Density: Four participants described seeking pages whose text was not visually "dense," i.e., pages with short sentences, short paragraphs, short line lengths, ample whitespace between lines, and/or using outline features such as bullet points, headers, or highlighting of key points.

The layout of the SERP itself was often confusing to users with dyslexia, as modern SERPs are often cluttered with links, advertisements, inline answers, right rail content, etc., and the density, spacing, colors, and font families used in the SERP are not typically adjustable by the end-user. Reading challenges made navigating the SERP challenging, and some participants reported using software (e.g., Dragon) to read results pages aloud to them; others wished for more multi-modal options in the SERP, such as the ability to highlight words and hear them read aloud, or hover over words and see an image representation of that term. Half of participants noted difficulties in recognizing domain names or URLs, and several suggested that iconic or screen-shot representations of pages on the SERP would help them identify familiar pages and/or identify key features (e.g., density, presence of multimedia) when choosing among unfamiliar pages.

Information Extraction

Upon selecting a target webpage from the SERP that met the criteria described in the prior section, our participants still described challenges in locating the sought piece of information within that page (a step that is necessary to have a successful information-seeking session [1]). Half of participants discussed a desire for voice output, and several had taken the step of installing software (e.g., Speaky, Dragon, Dream Reader) that could read web pages aloud to them for easier comprehension. P8 had been taught to highlight articles in different colors to support better reading comprehension, so she reported printing copies of web pages so she could engage with them in this manner.

Visual characteristics of webpages made reading more challenging for some participants, such as font type (several mentioned having trouble reading italics, several mentioned

preferring sans serif fonts), font size (with larger sizes facilitating reading), and the color and contrast of the page's visual scheme. Some of these presentation features could be controlled by the user if they were knowledgeable about how to use their browser and if the webpages were coded in a flexible way, while sometimes these features were beyond a user's control and impacted their ability to read and extract the target information.

Search Engines as a Linguistic Tool

Finally, our participants also described executing a number of queries whose goal was not to find a target website per se, but rather to answer a specific linguistic question (i.e., a specific type of transactional [7] rather than informational search). Fourney et al. [13] observed that language-related queries (e.g., seeking information related to spelling, definitions, or grammar) accounts for around 3% of search engine traffic; our interviews suggest this task may be much more common among information-seekers with dyslexia.

For instance, four of our participants described how reading out loud (i.e., pronouncing written words) was one of their challenges from dyslexia, and how they would use search engines to search unfamiliar words so that they could discover audio clips as pronunciation exemplars. Three participants described using search to help them understand the correct usage of words; for instance, P9 was confused about when to use the spelling *weather* versus *whether* and recalled searching queries such as "correct way to use weather" and "when to use whether or weather." Two reported searching for the etymology of words, as understanding etymology was a strategy they had been taught to improve their writing and reading skills. Two participants also mentioned using search engines' spell-checking features to support writing tasks in other applications/modalities; for instance, P1 noted she would query words to see their corrected spelling whilst creating handwritten documents.

ONLINE QUESTIONNAIRE

Our interview study revealed a set of rich vignettes detailing the challenges and coping strategies employed by dyslexic web searchers, and, to this end, made great strides in answering our first research question; however, one limitation of the interview method is that long interviews limit the sample size that can be reached. In order to validate that the themes from our interviews were relevant to a broader sample of users, we developed an online questionnaire. Further, to answer our second research question, we deployed the questionnaire both to adults with dyslexia and also to adults without dyslexia, so that we could compare and contrast the extent to which some of these themes resonated with each group.

Online Questionnaire: Method

We recruited adults who lived in the U.S. and were fluent in English to take our online questionnaire during a one-week period in August 2017. We advertised the questionnaire via social media, purchasing promoted posts on Twitter and Facebook, targeted toward people who either used the hashtag *#dyslexia* or who followed or liked organizations associated with dyslexia. To incentivize participation, our organization pledged to donate $1 to the International Dyslexia Association for each completed survey.

The survey consisted of 35 questions, and took about ten minutes to complete. There were a few demographic items (age, gender, level of education, dyslexia status), followed by multiple-choice questions about web search habits and skill and preferences for or challenges with search features that arose in our interview study. Lastly, there was a set of recent critical incident questions [11] in which we asked participants to reflect on a recent challenging search task (i.e., one that required more than 5 minutes to complete); we provided links to allow them to open their Google or Bing search histories in case they needed to refresh their memory. We asked them to tell us the goal they were trying to accomplish, the search query they initially tried, whether they were successful in their search, and to describe any challenges experienced during this specific search task.

Because of the ordinal nature of Likert-type responses, we use non-parametric statistical tests in our analyses. All significance values have been adjusted for multiple comparisons using Bonferroni corrections.

Online Questionnaire: Participants

We received 161 valid survey responses, 80 from adults who self-identified as having dyslexia, and 81 from adults who self-identified as non-dyslexic. Of the 80 respondents who identified as dyslexic, 50 indicated they had been professionally diagnosed and 30 indicated they were self-diagnosed. Mann-Whitney *U* tests comparing the professionally- and self-diagnosed participants' responses to all Likert-type questions reveal no statistically significant differences between these sub-groups on any items, so we group them both together into a single "with dyslexia" group for all of the analyses that follow. We refer to the survey respondents with identifiers R# (to avoid confusion with the earlier interview participants); id numbers go above 161 since the survey software also assigned identifiers to partially-complete responses.

146 participants (91%) identified as female; since the gender ratio for dyslexia is typically considered to be near parity [30] or even skewed slightly male [15], we surmise that women may be more likely to follow dyslexia-related interest groups on social media. Women are more likely to use Facebook in general [14], and comments in some of the responses from the non-dyslexic group suggest that many respondents from this category are mothers of children with dyslexia or special-education teachers (an occupation that skews highly female [9]). This is a limitation of using social media ads for recruiting for this population.

Participants' ages were distributed over a wide range: 18 – 24 (7.5%), 25 – 34 (8.1%), 35 – 44 (37.9%), 45 – 54 (28.6%), 55 – 65 (13.7%), 65 – 74 (3.7%), and 75+ (0.6%).

Participants' highest level of education included a range of levels: less than high school (1.9%), high school degree (3.7%), trade/technical school (3.7%), some college (11.2%), associate's degree (6.2%), bachelor's degree (32.3%), and graduate degree (41.0%). Transforming educational attainment into an ordinal scale, a Mann-Whitney *U* test finds that educational attainment differed significantly between the dyslexic and non-dyslexic groups, $z = -5.44$, $p < .001$, with the non-dyslexic users more likely to have attained the highest level on the scale (an advanced degree, median = 6) than those with dyslexia (median = 5, bachelor's degree); this may reflect the presence of special education professionals (who earn a Master's degree) in the participant pool, and may also reflect that the challenges associated with dyslexia may limit educational attainment.

Online Questionnaire: Findings

First, we examine the responses from the 80 participants with dyslexia, to examine whether the themes uncovered in our interview study resonate with this larger sample (RQ1). Then, we compare and contrast the questionnaire responses of participants with and without dyslexia (RQ2).

RQ1: Respondents with Dyslexia

Since query formulation challenges were a major theme in our interview study, our survey included several questions about strategies for entering search queries.

Like our interview participants, questionnaire respondents with dyslexia reported frequent use of voice input when performing web searches on mobile phones. When asked how often they use voice input to search the web on a mobile device with choices "Never" (0), "Monthly" (1), "Weekly" (2), or "Daily" (3), the median response was 2 ("weekly") and the modal response was "daily" (38.8% of respondents). When asked the same question about using voice input for search on a desktop or laptop computer, the modal response was "never" (58.8%), although 13.8% reported using voice to input search on their computer weekly and 12.5% daily. When asked about their preferred method to input search queries (regardless of device type), 61.3% chose "text", and 33.8% chose "voice". An additional 5% of respondents chose "other", then used the space provided to explain that their preference was highly context dependent. For example, one respondent preferred using a keyboard at the computer and voice with her phone (R34), echoing the trend outlined above. Two other respondents cited social reasons not to use voice in public, such as disturbing coworkers (R32) or looking silly (R88).

We also asked how respondents dealt with spelling errors when using a search engine. 66.3% indicated agreement that they "rely heavily on autocomplete" (where the search engine offers word and phrase completions within the search box itself as the query is being composed). 71.3% indicated using the search engine's "did you mean" feature (wherein the search engine suggests a new spelling after a query has been entered). 20.0% said that they compose their query in a separate text editor (so that they can use the text editor's spellchecker) and then copy and paste the corrected text into the search engine. 35.0% said they use the search engine as a linguistic tool by conducting an intermediate web search asking how to spell a certain word so that they can then use that information to compose their intended query. Additionally, 10.0% wrote in other methods they use to address correctly spelling their search query. Two mentioned using a thesaurus if other listed methods failed (R122, R293). One participant used Grammarly to check for mistakes (R274). R285 said she relies heavily on text-to-speech. R70 mentioned asking someone for help.

Our interview participants mentioned several ways in which visual properties of the SERP and individual websites impacted their ability to find information. The questionnaire included questions about page properties to gauge the generality of these preferences. 22.5% of dyslexic participants reported using ad blocker software to help declutter pages. 80.0% reported using techniques to increase the size of text on web pages. 41.3% increased the font size in their web browser. 56.3% zoomed in on websites to improve text readability. 5.0% mentioned using other techniques to enlarge text, such as using the browser's reader mode (R208), using a magnifying glass (R35), or switching to a laptop instead of a smartphone (R285).

Respondents used a five-point Likert scale (1 = strongly disagree, 5 = strongly agree) to rate their level of agreement with several statements about factors that influence their preference for web pages. 78.8% agreed or strongly agreed that they prefer pages that make use of pictures or videos rather than relying only on text (mean = 3.96, median = 4). 68.8% agreed or strongly agreed that they prefer web pages that use lists and tables rather than paragraphs to organize text (mean = 3.91, median = 4). 47.6% agreed or strongly agreed that they prefer web pages that use "simple, easy-to-read vocabulary" (mean = 3.45, median = 3). 27.6% said they prefer to browse the mobile versions of web pages rather than the desktop version (due to structural simplifications often applied to pages for use on smaller screens) (mean = 2.75, median = 3).

In an open-ended question, we asked participants to describe strategies they use to make reading websites easier. Fifteen (18.75%) mentioned having text read aloud, ranging from whole pages to specific words. Participants also frequently mentioned decluttering webpages (10%) by, e.g., using the browser's reader mode (5%), covering parts of the screen with their hands, or zooming in so only a small portion of the screen is shown. Seven participants (8.75%) mentioned printing off material to read, of which two mentioned highlighting with markers. An additional six (7.5%) mark their place while reading by highlighting the text with the mouse or following along with the cursor or a finger. Several mentioned skimming or looking for headings, lists, and underlined or bolded phrases to avoid having to read the whole page (6.25%). Others had strategies for the search itself, like searching for images instead of webpages (R178).

Finally, we asked respondents to describe in their own words what they find most challenging about searching for information online. Familiar themes emerged: Twenty-three respondents mentioned difficulties spelling keywords (28.75%), and nineteen reported challenges with identifying relevant search results (23.75%).

In summary, our questionnaire findings from 80 participants with dyslexia indicate that the themes we observed in interviews generally hold for a larger audience, including preferences for voice input for queries, the use of several strategies for spelling correction of queries (including the use of external text editors), a preference for large fonts, images, and lists over paragraphs. Even highly elaborate strategies, such as physically highlighting printouts of web pages, were reported by both the interviewees and by the questionnaire respondents. Together, the interviews and questionnaire responses address RQ1 by robustly characterizing the challenges people with dyslexia encounter when searching for information online, and the mitigation strategies that they employ. We now address RQ2 by comparing the responses of both dyslexic and non-dyslexic respondents.

RQ2: Comparing Dyslexic and Non-Dyslexic Respondents

It may be the case that many of the things people dyslexia find challenging about web search are also challenging to the general population. For this reason we also had 81 people without dyslexia complete the questionnaire. For questions with ordinal scales, we use Mann-Whitney *U* tests to compare the two groups, and for questions with binary response options we use Pearson Chi-Square tests to compare differences in proportions.

People with dyslexia were significantly more likely than those without dyslexia to prefer voice input over typing for entering search queries, $\chi^2(2, N = 161) = 11.98$, $p = .003$. There was also a trend toward people with dyslexia using voice input more frequently on mobile devices than people without dyslexia, but this difference was only marginally significant ($z = 1.81$, $p = .07$). Likewise, for the recent critical incident search task, there was a marginally significant trend toward people with dyslexia being more likely to have used voice input for that task ($\chi^2(2, N = 161) = 5.05$, $p = .08$). Finally, there was no significant difference between the groups in the frequency of using voice input to perform search on desktops or laptops. As noted earlier, use of voice search may depend not only on preference, but also on external factors such as social context, and the device's ability (real or perceived) to accept voice input.

People with dyslexia were significantly more likely to express a preference for web pages containing pictures and videos than people without dyslexia ($z = 2.75$, $p < .01$). People with dyslexia were also significantly more likely to prefer web pages that used tables or lists rather than paragraphs to organize information ($z = 3.30$, $p = .001$). There was no significant difference between groups in terms of their preference for web pages that use easy vocabulary, or in terms of their preference for using the mobile version of a website instead of the desktop version.

People with dyslexia were also significantly more likely than those without dyslexia to rely on techniques to correct search query spelling, including autocomplete ($\chi^2(1, N = 161) = 13.75$, $p < .001$) and conducting separate search engine queries about how to spell words ($\chi^2(1, N = 161) = 5.59$, $p = .02$); they were marginally more likely to use the spelling correction tools in a separate text editor and then paste the corrected words into the query box ($\chi^2(1, N = 161) = 3.25$, $p = .07$). There was no significant difference between each group's reported use of the "did you mean" spelling suggestions produced by the search engine.

Respondents with dyslexia were significantly more likely to increase the font size of text in their web browser, $\chi^2(1, N = 161) = 6.73$, $p = .009$. There was no difference between the two groups in the likelihood of zooming in on a web page to increase readability. There was no difference in the likelihood of using an ad blocker between the two groups.

Finally, respondents with dyslexia were also significantly more likely to use special hardware or software to simplify inputting search queries, $\chi^2(1, N = 161) = 4.81$, $p = .03$.

In summary, the questionnaire also included people without dyslexia, allowing us to address RQ2 by comparing dyslexic and non-dyslexic populations. We found that while most of the preferences and strategies we identified are more strongly associated with dyslexic users (e.g., voice search, a preference for pages that include images, etc.), some findings applied to both groups (e.g., the tendency to do more voice searches on mobile devices).

ONLINE STUDY: RELEVANCE AND READABILITY

To further explore our second research question, we conducted an online study with 346 participants (174 with dyslexia) to understand how people with and without dyslexia rate the readability and relevance of web pages in the context of a web search task, and how the properties of these pages relate to these judgments.

Method

From the aforementioned questionnaire data, we selected ten of the queries and associated search goals that respondents with dyslexia described when reflecting on their most recent challenging web search. When selecting queries, we aimed to choose a set that covered diverse information needs and topics, that consisted of informational rather than navigational [7] queries, that would be of general interest and understandable with general knowledge, and that would not reveal personal details about any study participants. Table 1 shows the set of ten queries and associated goals.

For each query, we used the Bing Search API to download the twenty top-ranked search results for each query; in cases where one of the top twenty search results was not a web page (e.g., a PDF document or other download) or pointed to the same website as a higher-ranked result, we eliminated

Table 1. We selected these ten informational queries and descriptions of their associated search intent from the questionnaire responses of people with dyslexia, and showed them to participants in our online study to provide context for rating the usability and utility of twenty web pages returned by the Bing API for these queries.

Query	Intent
dog breed best for children	choosing a breed of dog to get for a family with young kids
Dublin things to do	researching places to see while on vacation in Dublin
grease stains, hints	getting a grease stain out of a shirt
how to make a balloon display	finding out how to create a balloon display
International Space Station	seeking information to track the position of the International Space Station
"rationale" vs "rational"	figuring out which word, "rationale" or "rational", is right for the context in which you want to use it
sea star wasting disease	finding information about sea star wasting disease for a report for a marine biology class
video game testing certificate	figuring out how to get a professional certificate in video game testing
Westie skin issues	learning about health concerns associated with the West Highland Terrier dog breed
Yogi bear Hagerstown	looking up information about a campground you might visit

that result from the set and took the next-ranked results until we had twenty. This yielded a total of 200 web pages (10 queries * 20 results each). We recorded the relative rank (1-20) of each web page for that query, cached the source for the page for analysis, and created a high-quality screen capture of each web page.

Participants first answered a brief demographic questionnaire similar to that from our survey. Then, they were randomly assigned one of the ten search queries. The twenty screenshots of the pages pointed to by the search results for that query were shown to the participant one at a time, in a random order; the participant was not aware of the search engine ranking associated with any of the pages. Before seeing each screenshot, the participant was given the following instructions: "Imagine that you entered the query <query terms> into a search engine because you wanted to find out about <goal>. The search engine returned the following page to you; please look over the page with this scenario in mind. When you have scrolled to the bottom of the page, we will ask you a few questions about it."

There are no existing scales for measuring the accessibility of web pages to users with dyslexia, so we developed a set of questions based on the rubrics of WebQual [24], and the Marshal Readability Checklist [28], as well as based on the themes that emerged from our interview and questionnaire studies. All questions used a five-point Likert scale (1 = strongly disagree, 5 = strongly agree), and participants were asked to rate their level of agreement with the ten statements shown in Table 2. The study took about thirty minutes; participants received a $20 gratuity.

We use non-parametric statistical tests to analyze the rating results, due to the ordinal nature of Likert-type scales. In addition to analyzing results for each question individually, we combined our ten questions into two scores (Table 2). Our *readability score* combines the first seven questions; negatively-phrased questions' responses (Q2, Q4, Q6) are inverted by subtracting their value from 6 so that higher scores are always better, and so that each question would contribute between 1 to 5 points to the scale. As such, the full readability score ranges from seven to thirty-five. Our *relevance score* combines the last three questions, again inverting the response to negatively phrased item Q10, for a relevance score that ranges from three to fifteen. To check the validity of these scales, we computed full correlation matrices of the component questions (Appendix A).

Participants

We deployed our study during the first week of September 2017; we advertised via Facebook, targeting people interested in dyslexia. English-speaking adults in the U.S. who used a web search at least once per week were eligible.

346 participants (174 with dyslexia and 172 without dyslexia) rated at least one web page. These numbers do not include responses that we eliminated as spam, identified by unusually fast completion times. We had 279 complete responses in which participants rated all twenty results for their assigned query, and 67 partial responses. We include both in our analysis, yielding a total of 5,972 page-rating sets. Each of these 5,972 page-rating sets included answers to our ten Likert-type questions about page readability and relevance; of these 59,720 Likert responses, 42 were excluded from analysis due to a software error that failed to record these values, yielding a total of 59,678 ratings.

Of these 346 participants, 174 identified as dyslexic and 172 as non-dyslexic. Of the dyslexic participants, 139 contributed full data. For the non-dyslexic participants, 140 contributed full data. Of the participants with dyslexia, 36.2% indicated they were self-diagnosed and 63.8% had a formal diagnosis; the average readability and relevance index scores for these two sub-groups of dyslexic participants showed no statistically significant differences, so we group them together for our analysis.

60.7% of participants were female, 39.0% were male, and one participant indicated "other" as their gender identification. Age distributions were: 18 – 29 (30.6%), 30 – 39 (30.9%), 40 – 49 (24.6%), 50 – 59 (8.4%), 60 – 69 (4.3%), and 70 and older (1.2%). Participants' highest level of education were: less than high school (2.9%), graduated high school (26.6%), trade/technical school (4.9%), some college (19.7%), associate's degree (15.0%), bachelor's degree (9.8%), and advanced degree (21.1%).

Table 2. After viewing a screenshot of a web page returned by the Bing API for the given search query and intent, participants rated their agreement with these ten statements on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). The table is split in two, with the top and bottom sections indicating questions included in our readability and relevance scales, respectively. The polarity column shows whether a question's score adds directly to its corresponding scale (+) or is inverted by subtracting from 6 (-). We show the mean (μ) and median (M) scores for each question and each scale, computed from the 174 dyslexic and 172 non-dyslexic participants, from whom we collected a total of 5,972 sets of ratings distributed over 200 web pages (10 queries x 20 results). Mann-Whitney U tests find the differences between dyslexic and non-dyslexic responses to be highly statistically significant ($p << 0.001$) for Q1-Q8, Q10, and for both combined scales (i.e., readability and relevance). For Q9, we found $p = 0.01$.

ID	Statement	Polarity	μ (M) Dyslexic	μ (M) Non-dyslexic
Q1	major points were clearly stated	+	3.21 (3)	3.48 (4)
Q2	design choices made reading harder (fonts, colors, etc.)	-	2.89 (3)	2.65 (2)
Q3	images and videos helped convey the main ideas	+	2.80 (3)	3.01 (3)
Q4	it was easy for me to lose my place while reading	-	2.84 (3)	2.63 (2)
Q5	the information was well-organized	+	3.02 (3)	3.33 (4)
Q6	I was distracted by banners and advertisements	-	2.89 (3)	2.59 (2)
Q7	overall the website was easy to read	+	3.09 (3)	3.40 (4)
	Combined readability score: Q1 + Q3 + Q5 + Q7 + (6 - Q2) + (6 – Q4) + (6 – Q6)		21.50 (22)	23.25 (23)
Q8	the web page was relevant to the web search task	+	3.20 (3)	3.63 (4)
Q9	I was able to find all of the information necessary to accomplish the web search goal	+	3.03 (3)	3.10 (3)
Q10	I couldn't understand why the search engine would return this page for this search task	-	2.72 (3)	2.22 (2)
	Combined relevance score: Q8 + Q9 + (6 – Q10)		9.50 (10)	10.52 (11)

80.3% of respondents reported using a search engine daily, and the other 19.7% reported using a search engine at least once per week. Respondents self-rated their search engine skill level on a three-point Likert-type scale: 9.8% self-rated their search engine skills as novice (1), 60.4% as average (2), and 29.8% as expert (3). A Mann-Whitney U test indicates that people with dyslexia rated themselves as having lower levels of search engine skill than people without dyslexia (z = -4.50, $p < .001$). Non-dyslexic respondents self-rated as 3.5% novice, 58.1% average, and 38.4% expert, whereas dyslexic respondents were 16.1% novice, 62.6% average, and 21.3% expert (mean expertise 2.35 vs 2.05).

Results

The mean time spent rating each web page was 2.50 minutes (σ = 5.46) and the median was 1.29 minutes. Respondents with dyslexia spent a mean of 2.83 minutes (σ = 7.19) and median of 1.28 minutes rating each page, whereas those without dyslexia exhibited less variance in completion times, spending a mean of 2.27 minutes (σ = 2.79) and a median of 1.31 minutes. These completion time differences were not statistically significant.

Our analysis found statistically significant differences between the responses of dyslexic and non-dyslexic participants for each of the 10 questions (Table 2; in each case, $p \leq 0.01$ by Mann-Whitney U). For example, non-dyslexic respondents tended to agree more strongly that pages stated their points clearly, were well-organized, and were relevant to the query. Conversely, dyslexic respondents tended to agree more strongly that design choices made pages more difficult to read, that ads were distracting, and that they could not understand why the search engine returned the page for the query. We repeated the analysis stratifying the data by self-rated expertise, and found that these significant differences persisted for average and expert searchers, but not for the 9.8% of respondents who self-rated as novices, likely due to small n since we only had six novices in the non-dyslexic group.

We also consider the combined readability and relevance metrics, which aggregated scores across related questions. Here we found that readability and relevance are significantly correlated for both the dyslexic ($r_s = 0.33$, $p < .001$) and non-dyslexic ($r_s = 0.59$, $p < .001$) respondents, indicating that readability factors may influence relevance judgments regardless of dyslexia status.

In summary, this study expands our understanding of RQ2, but demonstrating systematic differences in how dyslexic and non-dyslexic users rate the relevance and readability of search results. We saw that dyslexic users self-rated themselves as having lower search skill, and consistently rated pages as less readable and as less relevant to the search task than their non-dyslexic counterparts (perhaps reflecting their ability to more easily find the query-relevant information within a page due to the readability advantage).

DISCUSSION

This paper is the first to present detailed findings about the use of general-purpose web search engines by English-speaking adults. Through interviews, an online questionnaire, and an online study, we gathered qualitative and quantitative data to answer two research questions:

RQ1: What challenges do people with dyslexia encounter when searching for information online, and what mitigation strategies do they employ?

RQ2: To what extent do the behaviors and preferences of dyslexic searchers differ from non-dyslexic searchers?

We found that searchers with dyslexia experience substantial challenges finding information with search engines. While aspects of page readability impact the perceptions of all searchers, we saw that dyslexic and non-dyslexic searchers have measurably different perceptions of page readability and relevance, and employ different strategies during the query, triage, and information extraction stages of search.

Design Implications

Our interview and questionnaire findings suggest there is substantial opportunity to improve web search for people with dyslexia. Since people with dyslexia form a sizable proportion of search engine users (e.g., up to 20% of English speakers [17]), such changes have the potential to improve the information-seeking process for a substantial number of people; indeed, our questionnaire and study findings suggest that many changes to improve usability for people with dyslexia may also benefit others, albeit to a lesser degree. Here, we describe design implications for search interfaces and ranking algorithms based on our participants' experiences. Some of these changes may benefit a broad audience, and come at relatively little cost (e.g., adding a speech input button next to the search box across all non-mobile platforms), and search engines may wish to provide them universally. Other features may only hold value for people with reading challenges (e.g., factoring reading level into ranking algorithms), and search engines may want to allow users to opt in (or out) of such features through menu options or the completion of personal profiles; alternatively, prior research has implicitly identified a number of health conditions through search engine interactions (e.g., cancer [32], pregnancy [12]), and automatic identification of searchers who are likely to be dyslexic may be possible and could be used to automatically enable valued interface features, though such an approach poses privacy concerns.

Implications for Search Interfaces

Our findings suggest that user interface changes on search engine websites and in web browser software have the potential to improve the accessibility of information seeking for people with dyslexia. The interfaces used for query entry and refinement could be improved by making it easy to input a voice search from all platforms and device types. Allowing users to toggle between a single choice for spelling auto-correction (the status quo) and a n-best list (an interface common in other software such as word processors) might be helpful, especially in cases where a spelling correction algorithm may not easily guess the word a searcher with dyslexia was aiming for as the top suggestion. Using spellcheckers trained specifically on the types of errors commonly made by users with dyslexia may also be valuable [36]. Redundant, multimodal cues can also help users be more confident that their query has been interpreted correctly, such as placing images or icons representing a word next to the query after it is input (or spell-corrected), and/or a button to play back an audio pronunciation of the query terms (recent work by Berget et al. [6] lends support to the idea that iconography may benefit users with dyslexia).

The SERP is the primary interface for triaging search results, and could be re-styled for users with dyslexia. Simplifying the page so that it is less dense may be helpful (e.g., showing fewer snippets, but perhaps with higher confidence if using new ranking techniques as proposed in the next section). Adding visual information such as webpage thumbnails [39] next to the SERP may also help searchers with dyslexia recognize previously-visited sites and/or determine if the visual properties of a site align with their reading abilities. More extensive inline answers for linguistic queries, such as those proposed by [13], may also benefit this audience.

Changes to web browsers could benefit not only triaging the SERP contents, but also locating the sought information within the ultimate target webpage. Adding targeted text-to-speech functionality into browsers would greatly improve accessibility; users could highlight individual words or passages to only have those read audibly. Allowing users to mark up a page via highlighting or underlining may also be valuable to people who are taught to employ active reading techniques on paper for enhanced comprehension. Being able to hover over any word within a webpage and see the top image search result associated with it would also improve the readability of many web pages for people with dyslexia.

Implications for Search Algorithms

While the specific details of search engines' ranking algorithms are proprietary, such systems are generally understood to determine the content and ordering of a SERP based on factors such as keyword matching (e.g., Okapi BM25 [38]), link structure (e.g., PageRank [31]), and learning from user interactions (e.g., clickthrough data [33]). Prior studies of web search behavior indicate that searchers typically click on the top few results, rarely exploring further in the results list, even though many results may be relevant [20]. However, our interview and study findings suggest that these ranking features may not place relevant pages that best match the needs of searchers with dyslexia near the top of the SERP. Our interviewees described that it was common for them to click back and forth between the results and the SERP, quickly until they found pages matching criteria that made them more suitable for consumption by people with dyslexia. Algorithms for ranking search results could take additional features into account that may be valued by searchers with dyslexia; such features might include aspects relating to reading level, page structure, and visual clarity.

Some reading level features have been proposed in the past for use in search ranking for children [8, 21, 23], by matching vocabulary words used within a webpage to lists of vocabulary associated with particular grade levels. Other possible approaches include computing statistics relating to word length, syllables-per-word, sentence length, grammatical complexity, and/or document length. Scores that build on these constructs include: Flesch Reading Ease, Flesch-Kincaid Grade Level, and Gunning Fog Index [29].

Aspects of page structure that could be automatically deduced from HTML could also feed into ranking

algorithms. For instance, scoring pages based on the presence and quantity of structural elements (e.g., headings, lists, tables) may offer value to people with dyslexia. Similarly, calculating a media-friendliness score based on the presence, amount, and distribution of imagery, audio, and/or video, and the ratio of multimedia to textual content may hold value. Similarly, some aspects of a page's visual clarity may be deduced from its HTML. While visual presentation features are sometimes used in ranking of results for search on mobile devices, we propose taking visual features into account on all platforms with respect to their impact on users with reading disabilities, i.e., scoring pages on factors such as the presence or absence of certain font families and styles, and the color scheme. For instance, the W3C contrast score [w3.org/TR/AERT#color-contrast] might be used to evaluate the readability impact of a page's foreground and background color palette. Alternatively, the scoring could take into account that some aspects of the page may be alterable by the end user and could score the page based on how customizable the appearance is, or on how well the page converts into standard high-contrast modes or reading modes offered by some operating systems and browsers. It may be that, because dyslexia is a spectrum disorder, it is important to personalize the weightings of readability features for a particular user's symptoms.

Some of these suggested changes to ranking algorithms could instead be manifested as user interface changes if the scores on these dimensions were shown directly to the end user in the SERP, allowing the user to choose which dimensions are of import to them (though this may make the SERP more complex to interpret). Understanding which of these ranking features may hold the highest value for searchers with reading disabilities, how to weight the scores for these features relative to traditional ranking factors, and how to best balance interface versus algorithmic enhancements is a key question for further study.

While we proposed several design implications for search interfaces and algorithms based on our interview findings, implementing these design ideas and verifying their efficacy for people with dyslexia is left to future work; exploring whether such interventions may also benefit other groups with language challenges (e.g., English language learners, children) is another open question. Similarly, while our investigation focused on web search, many of these design suggestions may also enhance the usability and inclusivity of other types of computing tasks and interfaces, particularly those for text-entry and text-consumption tasks (e.g., e-readers, e-textbooks, etc.).

Limitations

While our studies present initial insights on supporting information retrieval by people with language disabilities, it is worth bearing in mind the limitations of interview and survey studies, including possible inaccuracies or incompleteness in self-reported data, and the limited sample sizes that can be obtained through time-intensive interviews and questionnaires collecting lengthy open-ended responses. Since quick diagnostic tests of dyslexia (e.g., [34]) are not yet clinically accepted or available for the English language, we had to rely on respondents' self-report as to whether or not they have dyslexia. Additional studies of the impact of dyslexia on information seeking that use methods such as lab observations, log analysis, and/or large-scale surveys would offer valuable expansion of these findings.

The interplay of education level and dyslexia status is complex, since dyslexia may impact educational attainment, which may in turn impact search proficiency. In our survey data, we found that educational attainment had no statistically significant effect on dyslexic participants' reported search mitigation strategies. In our online study, we found that people with dyslexia self-rated as having lower search expertise than non-dyslexics, but that our findings held when controlling for search skill levels (see Appendix B). However, further study of the relationship among dyslexia status, educational attainment, and search skills is an important area for future work.

Self-selection bias may also impact our findings – it may be the case that people with more severe dyslexia-related challenges chose to respond to our call for participants in order to share their experiences; employing varied sampling methods in future studies can help give a better idea of the pervasiveness of specific search-related challenges in the population. Our non-dyslexic samples may also be biased by our recruitment method, in that our advertisement was targeted to people who indicated dyslexia or dyslexia-related organizations as an interest on social media, so likely includes friends, relatives, and teachers of people with dyslexia, who may somehow systematically differ from the broader population in ways we are unaware of.

CONCLUSION

In this paper, we identified that web search accessibility for people with dyslexia is an important and largely unaddressed challenge for the HCI and Information Retrieval communities. By conducting and analyzing interviews with ten adults with dyslexia as well as an online survey with 80 dyslexic and 81 non-dyslexic adults, we found that people with dyslexia experience challenges in finding information through web search relating to all stages of the information seeking process. We also conducted an online study in which 174 dyslexic and 172 non-dyslexic adults rated the readability and relevance of search results. From these studies, we identified several aspects of query formulation, results triage, and information extraction that are challenging for people with dyslexia, and found marked differences in dyslexic and non-dyslexic adults' search behaviors. Reflecting on these findings, we proposed several design suggestions for improving both search interfaces and search algorithms. These findings and design implications open new avenues for further research on improving the accessibility of information seeking and, more generally, the web, for people with reading disabilities.

ACKNOWLEDGMENTS

We would like to thank Shane Williams, Ryen White, and Kevin Larson for helpful "hacks" and discussions that contributed to our thinking on this topic.

REFERENCES

1. Ageev, M., Go, Q., Lagun, D., and Agichtein, E. Find it if you can: A game for modeling different types of web search success using interaction data. *Proceedings of SIGIR 2011*.
2. Al-Wabil, A., Zaphiris, P., and Wilson, S. Web Navigation for Individuals with Dyslexia: An Exploratory Study. *Universal Access in Human Computer Interaction 2007*.
3. Al-Wabil, A., Zaphiris, P., and Wilson, S. Examining visual attention of dyslexics on web navigation with eye tracking. *Innovations in Information Technology 2008*.
4. Berget, G. Search and find? An accessibility study of dyslexia and information retrieval. Doctoral Dissertation, University of Oslo, 2016.
5. Berget, G. and Sandnes, F.E. (2015) Searching databases without query-building-aids: Implications for dyslexic users. *Information Research*, 20(4).
6. Berget, G., Mulvey, F., and Sandnes, F.E. (2016) Is visual content in textual search interfaces beneficial to dyslexic users? *International Journal of Human-Computer Studies*. August 2016, Vol. 92, Issue C, pp. 17-29.
7. Broder, A. A Taxonomy of Web Search. *SIGIR Forum 2002*.
8. Collins-Thompson, K., Bennet, P.N., White, R.W., de la Chica, S., and Sontag, D. Personalizing Web Search Results by Reading Level. *Proceedings of CIKM 2011*.
9. DATAUSA: Special Education Teachers. https://datausa.io/profile/soc/252050/#demographics Retrieved August 25, 2017.
10. Figueredo de Santana, V., de Oliveira, R., Almeida, L.D.A., and Baranauskas, M.C.C. Web accessibility and people with dyslexia: a survey on techniques and guidelines. *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility 2012*.
11. Flanagan, J. The critical incident technique. *Psychological Bulletin*, 51:327-358, 1954.
12. Fourney, A., White, R., and Horvitz, E. Exploring Time-Dependent Concerns about Pregnancy and Childbirth from Search Logs. *Proceedings of CHI 2015.*
13. Fourney, A., White, R., and Morris, M.R. Web Search as a Linguistic Tool. *Proceedings of WWW 2017.*
14. Greenwood, S., Perrin, A., and Duggan, M. Social Media Update 2016. *Pew Research Center for Internet and Technology*. November 11, 2016. http://www.pewinternet.org/2016/11/11/social-media-update-2016/
15. Hawke, J.L., Olson, R.K., Willcutt, E.G., Wadsworth, S.J., and DeFries, J.C. (2009) Gender ratios for reading difficulties. *Dyslexia*, 2009 August, 15(3), 239-242.
16. Hoeft, F., McCardle, P., and Pugh, K. "The Myths and Truths of Dyslexia in Different Writing Systems." International Dyslexia Association blog, March 2015.
17. International Dyslexia Association. "Dyslexia Basics." https://dyslexiaida.org/dyslexia-basics/ Retrieved 11/23/16.
18. International Dyslexia Association. "Testing and Evaluation." https://dyslexiaida.org/testing-and-evaluation/ Retrieved 12/13/17.
19. Ivory, M., Sinha, R.R., and Hearst, M.A. (2000) Preliminary Findings on Quantitative Measures for Distinguishing Highly Rated Information-Centric Web Pages. *Proceedings of the Sixth Conference on Human Factors and the Web*.
20. Joachims, T., Granka. L., Pan, B., Hembrooke, H., and Gay, G. Accurately interpreting clickthrough data as implicit feedback. *Proceedings of SIGIR 2005*.
21. Kim, J., Collins-Thompson, K., Bennett, P.N., and Dumais, S. Characterizing Web content, User Interests, and Search Behavior by Reading Level and Topic. *Proceedings of WSDM 2012*.
22. Lazar, J., Feng, J.H., and Hochheiser, H. Research Methods in Human-Computer Interaction. *Wiley*, 2010.
23. Liu, X., Croft, W.B., Oh, P. and Hart, D. Automatic recognition of reading levels from user queries. *Proceedings of SIGIR 2004*.
24. Loiacono, E.T., Watson, R.T., and Dale L. Goodhue, D.L. (2007) WebQual: An Instrument for Consumer Evaluation of Web Sites. *International Journal of Electronic Commerce*, 11(4).
25. MacFarlane, A., Al-Wabil, A., Marshall, C.R., Albrair, A., Jones, S.A., and Zaphiris, P. (2010) The effect of dyslexia on information retrieval: a pilot study. *Journal of Documentation*, 66(3), 307-326.
26. MacFarlane, A., Albraier, A., Marshall, C.R., and Buchanan, G. Phonological working memory impacts on information searching: an investigation of dyslexia. *Proc. of the Information Interaction in Context Symposium*, 2012.
27. MacFarlane, A., Buchanan, G., Al-Wabil, A., Andrienko, G., and Andrienko, N. Visual Analysis of Dyslexia on Search. *CHIIR 2017*.
28. Marshall, N. (1979) Readability and Comprehensibility. *Journal of Reading*, 22(6), 542-544.

29. McCallum, D.R. and Peterson, J.L. (1982) Computer-based readability indexes. *Proceedings of the ACM '82 Conference.*
30. Miles, T.R., Haslum, M.N., and Wheeler, T.J. Gender Ratio in Dyslexia. *Annals of Dyslexia*, December 1998, Volume 48, Issue 1, pp. 27 – 55.
31. Page, L., Brin, S., Motwani, R., and Winograd, T. (1999) The PageRank Citation Ranking: Bringing Order to the Web. Technical Report. Stanford InfoLab.
32. Paparrizos, J., White, R., and Horvitz, E. Screening for Pancreatic Adenocarcinoma Using Signals from Web Search Logs: Feasibility Study and Results. *Journal of Oncology Practice*, 12(8), August 2016, p. 737-744.
33. Radlinski, F., and Joachims, T. Query chains: learning to rank from implicit feedback. *Proceedings of KDD 2005*.
34. Rello, L., Ballesteros, M., Ali, A., Serra, M., Sánchez, D.A., and Bigham, J.P. Dytective: Detecting Risk of Dyslexia with a Game. *Proceedings of Pervasive Health 2016*.
35. Rello, L., Pielot, M., and Marcos, M-C. Make It Big! The Effect of Font Size and Line Spacing on Online Readability. *Proceedings of CHI 2016*.
36. Rello, L., Ballesteros, M., and Bigham, J.P. A Spellchecker for Dyslexia. *Proceedings of ASSETS 2015*.
37. Rello, L. Dyslexia and Web Accessibility: Synergies and Challenges. *Proceedings of W4A 2015*.
38. Spärck Jones, K., Walker, S., and Robertson, S. E. A probabilistic model of information retrieval: Development and comparative experiments. *Journal of Information Processing and Management*, 36(6), November 2000, p. 779-808.
39. Woodruff, A., Faulring, A., Rosenholtz, R.E., Morrison, J., and Pirolli, P.L. Using Thumbnails to Search the Web. *Proceedings of CHI 2001*.

APPENDIX A: CORRELATIONS AMONG SCALE ITEMS

This appendix provides the correlation coefficients (Spearman's rho for non-parametric correlations) for the component questions of our readability and relevance scales. Correlations in bold are significant at the .01 level (two-tailed), and those in italics are significant at the .05 level (two tailed). We break out the scores separately for the 174 dyslexic and 172 non-dyslexic participants. Correlations were stronger for all scale items for non-dyslexic participants, which may reflect reading comprehension issues by participants with dyslexia, since the scales were administered in an online questionnaire.

Relevance (Dyslexic)	**Q8**	**Q9**	**Q10**
Q8	1.00	**0.57**	**-0.52**
Q9	**0.57**	1.00	**-0.42**
Q10	**-0.52**	**-0.42**	1.00

Relevance (Non-Dyslexic)	**Q8**	**Q9**	**Q10**
Q8	1.00	**0.75**	**-0.83**
Q9	**0.75**	1.00	**-0.72**
Q10	**-0.83**	**-0.72**	1.00

Readability (Dyslexic)	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**
Q1	1.00	*-0.17*	0.11	**-0.35**	**0.32**	-0.12	**0.40**
Q2	*-0.17*	1.00	0.13	**0.29**	-0.10	**-0.21**	**-0.25**
Q3	0.11	0.13	1.00	0.06	**0.34**	**-0.26**	*0.17*
Q4	**-0.35**	**0.29**	0.06	1.00	0.01	0.02	**-0.39**
Q5	**0.32**	-0.10	**0.34**	0.01	1.00	*-0.14*	**0.31**
Q6	-0.12	-0.21	**-0.26**	0.02	*-0.14*	1.00	**-0.20**
Q7	**0.40**	-0.25	*0.17*	**-0.39**	**0.31**	**-0.20**	1.00

Readability (Non-Dyslexic)	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**
Q1	1.00	**-0.39**	**0.23**	**-0.36**	**0.69**	**-0.21**	**0.55**
Q2	**-0.39**	1.00	**-0.27**	**0.55**	**-0.55**	**0.21**	**-0.65**
Q3	**0.23**	**-0.27**	1.00	**-0.23**	**0.23**	-0.11	**0.28**
Q4	**-0.36**	**0.55**	**-0.23**	1.00	**-0.45**	*0.16*	**-0.56**
Q5	**0.69**	**-0.55**	**0.23**	**-0.45**	1.00	**-0.28**	**0.62**
Q6	**-0.21**	**0.21**	-0.11	*0.16*	**-0.28**	1.00	**-0.34**
Q7	**0.55**	**-0.65**	**0.28**	**-0.56**	**0.62**	**-0.34**	1.00

APPENDIX B: SEARCH EXPERTISE ANALYSIS

Dyslexic participants self-rated as having lower average search expertise than non-dyslexics. It is unclear if this reflects a true difference in search expertise, or only differences in how users perceive their skill levels. It is also unclear if this self-rating difference is associated with having dyslexia, or reflects an artifact of how participants were recruited (or some combination of both). For example, it is possible that non-dyslexic individuals responding to the study advertisement were more likely to be healthcare or education professionals, and that these occupations impact their familiarity with web search. To investigate whether our findings might be attributable to differences in search skill, we repeated the analysis discussed in the Results section, but stratified by self-rated search expertise. In each strata, we found similar trends as shown in the Results section's Table 2, but the magnitude and significance of differences varied: Among experts, all differences were highly significant. Among those of average expertise, all differences were again highly significant except for Q3 and Q9. Finally, among novices, no significant differences were found, except for Q5. These findings indicate that the trends discussed in the paper hold even when adjusting for self-rated search expertise for the searchers with average or above-average skill, who constituted more than ninety percent of our participants (the non-significance of the findings for novice searchers is likely due to small *n*, since there were only six non-dyslexic searchers who self-identified as novice). Future work investigating the interplay of dyslexia, search expertise, and perceptions of that expertise would add to our understanding of issues facing searchers with dyslexia.

Table B1. We show the mean (μ) and median scores (M) for each question and each scale, computed from the 28 dyslexic and 6 non-dyslexic participants who self-rated their search expertise as *novice*. The 4th column shows the *p*-value computed with a Mann-Whitney *U* test, with values in bold indicating statistical significance.

PARTICIPANTS SELF-RATING AS NOVICE

ID	μ (M) Dyslexic	μ (M) Non-dyslexic	*p*-value
Q1	3.26 (3)	3.51 (4)	0.078
Q2	2.38 (2)	2.49 (2)	0.391
Q3	2.41 (2)	2.65 (2)	0.051
Q4	2.50 (2)	2.65 (2.5)	0.241
Q5	2.92 (3)	3.3 (4)	**0.003**
Q6	3.02 (3)	3.06 (3)	0.752
Q7	3.12 (3)	3.29 (3)	0.177
Readability	21.82 (22)	22.55 (22)	0.165
Q8	2.95 (3)	2.93 (3)	0.552
Q9	2.85 (3)	2.925 (3)	0.673
Q10	2.79 (3)	2.83 (3)	0.670
Relevance	9.01 (9)	9.025 (9)	0.465

Table B2. We show the mean (μ) and median scores (M) for each question and each scale, computed from the 109 dyslexic and 100 non-dyslexic participants who self-rated their search expertise as *average*. The 4th column shows the *p*-value computed with a Mann-Whitney *U* test, with values in bold indicating statistical significance.

PARTICIPANTS SELF-RATING AS AVERAGE

ID	μ (M) Dyslexic	μ (M) Non-dyslexic	*p*-value
Q1	3.19 (3)	3.47 (4)	**<< 0.0001**
Q2	3.00 (3)	2.67 (2)	**<< 0.0001**
Q3	2.90 (3)	2.96 (3)	0.132
Q4	2.89 (3)	2.66 (2)	**<< 0.0001**
Q5	3.01 (3)	3.32 (4)	**<< 0.0001**
Q6	2.89 (3)	2.60 (2)	**<< 0.0001**
Q7	3.07 (3)	3.38 (4)	**<< 0.0001**
Readability	21.40 (21)	23.20 (23)	**<< 0.0001**
Q8	3.23 (3)	3.61 (4)	**<< 0.0001**
Q9	3.09 (3)	3.08 (3)	0.964
Q10	2.72 (3)	2.25 (2)	**<< 0.0001**
Relevance	9.60 (10)	10.45 (11)	**<< 0.0001**

Table B3. We show the mean (μ) and median scores (M) for each question and each scale, computed from the 37 dyslexic and 66 non-dyslexic participants who self-rated their search expertise as *expert*. The 4th column shows the *p*-value computed with a Mann-Whitney *U* test, with values in bold indicating statistical significance.

PARTICIPANTS SELF-RATING AS EXPERT

ID	μ (M) Dyslexic	μ (M) Non-dyslexic	*p*-value
Q1	3.23 (3)	3.49 (4)	**<< 0.0001**
Q2	3.01 (3)	2.63 (2)	**<< 0.0001**
Q3	2.84 (3)	3.12 (3)	**<< 0.0001**
Q4	3.02 (3)	2.57 (2)	**<< 0.0001**
Q5	3.11 (3)	3.34 (4)	**<< 0.0001**
Q6	2.79 (3)	2.53 (2)	**<< 0.0001**
Q7	3.13 (3)	3.45 (4)	**<< 0.0001**
Readability	21.47 (22)	23.67 (24)	**<< 0.0001**
Q8	3.33 (4)	3.74 (4)	**<< 0.0001**
Q9	3.00 (3)	3.16 (3)	**0.014**
Q10	2.68 (2)	2.10 (2)	**<< 0.0001**
Relevance	9.65 (10)	10.80 (11)	**<< 0.0001**