Reading is Believing?

A Study of the Relationship Between Website Readability and Online Credibility

David T. Cormack

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

Master of Arts

Department of Psychology

University of Manitoba

Winnipeg

Abstract

More and more parents are turning to the Internet for information about their children's health and wellbeing. With all of the conflicting information online, how do parents decide which health websites to believe and follow? One website element the impact of which is not well understood is the readability level of content presented. An online survey was conducted with parents and university students in an attempt to isolate the relationships between website readability and credibility judgements and intentions to act. Parents and students were randomly assigned to view and evaluate one of three mock websites about Fragile X syndrome with identical content but differing readability levels. It was predicted that users would find websites that were easier to understand more credible and they would in turn then be more likely to act on the information from simple readability websites. Contrary to the author's hypothesis, students were more likely to follow recommendations on websites written at a complex readability level, when they were engaged with the website. Results also showed that readability was less of an influence on credibility than were the parents' pre-experiment knowledge of the disorder.

Keywords: online health information, readability, online credibility, involvement

Acknowledgements

This manuscript was completed under the supervision of Dr. Barry Mallin in partial fulfillment for the requirements of the M.A. in School Psychology at the University of Manitoba.

I would like to thank my advisor, Dr. Barry Mallin for his help and advice throughout this study, as well as acknowledge the expertise and support from my committee members, Dr. Randy Jamieson and Dr. Andrea Bunt. I would also like to acknowledge my collaborators, Erin Braun-Janzen, Hannah Burkholder, and Jon Canzio and thank them for their time.

I would also like to acknowledge the generous funding and support provided by the Canadian Institute of Health Research, the Faculty of Graduate Studies, and the Department of Psychology.

Correspondence concerning this manuscript should be addressed to David Cormack, Department of Psychology, University of Manitoba, Winnipeg, MB, R3T 2N2. E-mail: umcorma2@myumanitoba.ca

Table of Contents

List of Tables

List of Figures

Reading is Believing?

A Study of the Relationship Between Website Readability and Online Credibility

Since the emergent popularity of patient-centred treatment in the early 1990's (Mead & Bower, 2000), individuals today have more input into how their health is managed. This approach emphasizes a shared responsibility for decision-making between health professionals and patients, rather than a unilateral expert decision-making model (Mead & Bower, 2000). When the patient is a child, the parent takes on the role of collaborating with health professionals on behalf of the child. In order for parents to make informed health decisions they must have some understanding of their child's diagnosis as well as the treatment options available.

Since health professionals lack the time for extensive parent education the Internet has become a very popular health resource. There is now a vast ocean of online health information available to parents, which they can access on their own time and from the comfort of their home. It is estimated that 79% of Internet users in the United States (Fox, 2007) have accessed health information online. A Canadian study from 2005 reported that 8.7 million Canadian adults, 58% of Internet users, accessed online health information (Underhill & McKeown, 2008). Young parents are especially likely to seek out health information online, and the number of parents seeking health websites appears to be growing (Plantin & Daneback, 2009). Parents of young children seek information about childhood disorders (Plantin & Daneback, 2009) and value the Internet as a source of information (Turner, Kabashi, Guthrie, Burket & Turner, 2011). Users report that one reason for seeking online health information is to take control of their health (Sillence, Briggs, Harris & Fishwick, 2006).

Ideally the accessibility of health information should empower parents to better care for their children and lead to more productive collaborations between parents and professionals.

However, the Internet is unregulated, leading to conflicting views that often engender confusion rather than clarity. A growing body of research agrees that the quality of health websites is unpredictably varied (Breckon, Jones, Morris, & Richardson, 2008; Eysenbach, Powell, Kuss, & Sa, 2002; Reynolds, Walker, Walsh, & The Mobilizing Minds Research Group, 2014) and the information presented is often incomplete or misleading (Akram, Boyter, & Thomson, 2010; Burneo, 2006; Eysenbach, et al., 2002; Mitchell & Read, 2011; Murphy, Frost, Webster, & Schmidt, 2004; Oermann, Lowery, & Thornley, 2003). Parents must sift through the online health information available and make judgements about which websites to trust and follow.

With the varying quality of information contained on the Internet, online credibility has emerged as an important area of research, particularly in the area of computer science (Vega, Montague & DeHart, 2010). Definitions of online credibility are varied, with little agreement among researchers working on the problem (Vega, Montague & DeHart, 2010). In its simplest form, Fogg et al. (2001) equates credibility with the believability of a website. In a more systematic approach to defining credibility, Hilligoss & Rieh (2008) interviewed students and had them fill out detailed journals concerning their web search behaviours. The researchers then performed a content analysis of these in-depth search records, the results informing the development of an extensive model of credibility assessment. Their definition of credibility includes five characteristics: truthfulness, believability, trustworthiness, objectivity, and reliability. An empirically tested five-item Likert scale has been developed based on this definition (Lederman, Fan, Smith & Chang, 2014). Schwarz and Morris (2011) synthesized much of the existing literature, including the definitions created by both Fogg et al. (2001) and Hilligoss and Rieh (2008) to create an operational definition of credibility:

> A credible webpage is one whose information one can accept as the truth without needing to look elsewhere. If one can accept information on a page as true at face value, then the page is credible; if one needs to go elsewhere to check the validity of the information on the page, then it is less credible. (p. 1247)

For the purposes of this study, the definition presented by Hilligoss & Rieh (2008) will be used as the primary definition of credibility due to the breadth of concepts included within it and in order to leverage the strong psychometrics of the scale developed by Lederman, Fan, Smith and Chang (2014).

Website credibility is important to understand because users who trust health websites are likely to act on the information presented. Mou and Cogen (2014) found a significant positive correlation between website trust and intentions to act, $r = .263$, $p < .01$. Intentions to act have been described as the "best and most proximal psychological predictor of actual behavior" (Hu & Sundar, 2010). If parents believe the information they are finding online, they are likely to act, which may include following the site's recommendations, discussing the site with a medical professional (Sillence, Briggs, Harris & Fishwick, 2007), or sharing that information with others (Hu & Sundar, 2010).

In order to decide whether a website is worth trusting and acting on, parents must be able to understand the information and recommendations the website presents. The average reading level of Canadians is about an 8th grade level (McInnes & Haglund, 2011). However, 48.5% of Canadians are considered to have low literacy skills (Statistics Canada, Employment and Social Development Canada, and Council of Ministers of Education, Canada, 2013) and 55% of Canadian adults are considered to have low health literacy skills (Canadian Literacy and Learning Network, 2012). Therefore, in order for the majority of Canadians to comprehend

online health information, it must be written at a low level of complexity. Readability is a measure of the complexity of written text, and readability grade level scores estimate the number of years of education required in order to comprehend a written passage. For instance, a reading level score of 6 means that a minimum of a grade 6 education is recommended in order to fully understand a passage of text and reading level scores above grade 12 suggest the need for post-secondary education. The higher the readability level of a passage of text, the more difficult it is to comprehend.

Despite the low levels of literacy among the general population, studies have shown that online health information typically receives high readability grade levels (Brigo, Otte, Igwe, Tezzon & Nardone, 2015; Cheng & Dunn, 2015; Cormack, 2013; Huang, Fang, Agarwal, Bhagat, Eloy & Langer, 2015; McInnes & Haglund, 2011; Patel, Sanghvi, Cherla, Baredes & Eloy, 2015; Sobota & Ozakinci, 2015). In a comprehensive study, McInnes and Haglund (2011) examined 352 health websites covering 22 different health concerns and found an average readability score above a grade 12 level, with no websites written below a grade 7 level. This pattern of high readability applies to the realm of mental health websites as well (Reynolds, Walker, Walsh, & The Mobilizing Minds Research Group, 2014). As part of a previous study, the present author examined websites about attention-deficit hyperactivity disorder (ADHD), including their reading level. It was found that the readability score of the 39 websites selected ranged from grade 6 to 14 (post-secondary), with an average above grade 11 (Cormack, 2013). These studies assume that lower levels of readability will be beneficial to health consumers, such as parents. However, there is heavy criticism over the validity of readability measures (Redish, 2000; Schriver, 2000). There is also debate as to which of the many existing readability formulae are most effective (Begeny & Greene, 2014).

Readability measures have several limitations. Firstly, they are overly formulaic. The Flesch-Kincaid Grade Level and Flesch Reading Ease are among the most popular readability measures due in large part to their accessibility; each are included in Microsoft Word and can be automatically calculated whenever a spelling and grammar check is completed. Both Flesch reading measures are calculated based on the length of syllables, words and sentences within the selected passage of text. The formulae can be found in Table 1, along with that of other popular readability measures. The benefit of using readability formulae is that they can be calculated quickly and conveniently, producing a single number making interpretation of the score easy and intuitive. The downside is that formulae can only measure things that are countable (Redish, 2000), ignoring visual elements such as tables, chart, graphs, and images which all contribute to understanding (Schriver, 2000). Readability formulae also fail to consider text features that are difficult to quantify, such as writing style, mood, tone, and even word order (Redish, 2000). These formulae also fail to incorporate the benefits of layouts that improve ease of reading, such as bulleted lists and white space (Redish, 2000). Because readability scores require periods to know when a sentence ends, elements such as tables and bulleted lists actually increase the measured difficulty of text because the formula interprets them as one large sentence (Redish, 2000; Schriver, 2000). Their inability to incorporate graphic information and visual design make them particularly inappropriate for website evaluation. As one critic points out,

> As we come to understand the need to design information more visually on the web, readability formulas will have even less value for judging the usability of web pages. A good webpage might have not a single complete prose sentence – and that would make a readability formula totally inappropriate as a measure of either readability or usability for web sites. (Redish, 2000, p. 135)

The most fundamental limitation of readability measures is a lack of face validity, the discrepancy between what users want to know and what readability really means. Readability grade levels are used to assess how well an audience will be able to comprehend the message from a passage of text, but readability is only one small component of the process of comprehending. Comprehension is an interactive process that involves characteristics of the reader, such as background information, context, and engagement. Readability is an estimate of the complexity of text, which does not induce comprehension in readers. Without taking into account characteristics of the reader or the context of the information presented, readability can only partially predict comprehension.

Despite the criticisms concerning readability so clearly presented by Redish and Schriver 15 years ago, readability of online health information, primarily using the Flesch measures available in Microsoft Word, continues to be a popular area of study (Brigo, Otte, Igwe, Tezzon & Nardone, 2015; Cheng & Dunn, 2015; Huang, Fang, Agarwal, Bhagat, Eloy & Langer, 2015; Patel, Sanghvi, Cherla, Baredes & Eloy, 2015; Sobota & Ozakinci, 2015). A search on the medical research database Pubmed for "flesch" produces over 1,000 results, with the number of studies published each year continuing on an upward trend (see Figure 1). This suggests that researchers continue to view readability measures as valuable tools. Readability measures are widely available, easy to compute and interpret, and are correlated with comprehension (Redish, 2000).

The fundamental question of interest in the present study is: what is the relationship between website readability and online credibility judgements? A few different theories of online credibility have been proposed, each of which make unique predictions for how readability impacts these kinds of judgements.

Prominence Interpretation Theory (Fogg, 2003) suggests that readability has very little impact on website credibility. According to Fogg's theory, website credibility is a function of the judgements made about the features of the website. Judgements of website features are a two-step process: first, users must notice a feature (prominence), and secondly they must make some form of judgement (interpretation) based on the feature (Fogg et al., 2003). Features that are more prominent are more readily noticed and thus exert greater influence. This theory stemmed from a large-scale online study with 2,500 participants. Subjects viewed randomly selected websites, judged their credibility, and gave comments on how those judgements were made. Almost half, 46.1%, of all comments mentioned the visual design of the sites, while information clarity, which follows the definition of readability used in the present study, appeared in only 3.7% of all comments (Fogg et al., 2003). The theory suggests that since the complexity of text is not a feature that people readily recognize at first glance, people do not take it into account when making credibility judgements about a website. Prominence Interpretation Theory has lead other researchers to create ways of making website features that signal credibility, such as expert certifications and author credentials, more visible and prominent to users (Schwarz & Morris, 2011). However, displaying many features can create visual clutter while taking away from the prominence of other features (Olteanu, Peshterliev, Liu & Aberer, 2013).

A competing theory makes the case that the prominence of a feature is not always the main factor users consider when judging the credibility of websites. Sillence, Briggs, Fishwick and Harris (2007) propose a staged-model of trust, wherein users first make rapid judgements based on negative website features to eliminate websites that appear untrustworthy. Building trust in a website is an active process that entails a more systematic evaluation of the positive

characteristics of a website. This theory was tested with a group of 15 women undergoing an actual health concern. These women took part in web searches, group discussions, interviews and kept journals of their online search behaviours for a period of six months. While visual design factors were most important during the initial rejection of websites, content factors, including readability, were cited more often as part of developing trust in a website. The authors emphasized the importance of engagement on the part of their subjects. Because these women were searching for answers about a real health condition, it was suggested that they were better able to engage in the second stage of creating trust, the systematic evaluation that was found to be based more on content than design. These results suggest that readability does have a place in the evaluation of online credibility.

Ferebee (2007) further studied the role of user engagement on credibility judgements. She based her work on Prominence-Interpretation theory and provided evidence suggesting that different types of engagement caused users to notice different kinds of website features. Engagement was measured across two dimensions: enduring involvement and situational involvement. Enduring involvement is persistent and intrinsic to the individual, while situational involvement can be induced and is temporary (Ferebee, 2007). In her 2x2 design, Ferebee had users view and provide credibility judgements for websites about financial investments. Financial planners and university students were recruited to provide the enduring and non-enduring involvement groups respectively. Participants in the situational involvement groups were informed that they would be expected to answer questions after viewing the site. The groups who were induced into high situational involvement tended to notice content factors, including readability, more than the groups who were not induced into situational involvement. They also noticed more content factors than design factors. These results are consistent with the

work of Sillence, Briggs, Fishwick and Harris (2007) in showing that website content factors are valued when users are engaged with the websites they are judging.

A common element of the work outlined above is a reliance on self-report measures to validate the proposed theories. The researchers only considered website features if users mentioned those features in comments explaining how they arrived at their credibility judgements. It is possible that self-report measures are likely to be biased towards prominent features. Participants are only able to report features that they are aware of, and which are therefore prominent. Website features that are impactful, but less prominent, such as readability, may not be properly represented in self-report measures. Also, readability was not reported in isolation as a factor in most of the above studies, but was considered along with other content features. In order to fully investigate the relationship between readability and online credibility other research methods are required.

A group of researchers from Switzerland set out to predict website credibility automatically from known website features using computer modelling techniques (Olteanu, Peshterliev, Liu & Aberer, 2013). They examined a large database of websites that had been previously given credibility ratings by researchers with input from experts (Schwarz & Morris, 2011), using computer modelling to determine which website features best predicted the credibility ratings. Readability scores were among the best predictors of website credibility, ranking among the top five most predictive among 37 website features. Credibility and readability were positively correlated, such that websites rated as more credible tended to have higher levels of readability. The direction of this relationship runs counter to the assumption of many of the previously cited studies of readability among health websites, which is that lower readability levels are preferred when promoting health information online in order to promote

comprehension. This study also suggests that while readability measures may have flaws and users may not always recognize the complexity of text, they can, nevertheless, be highly correlated with credibility judgements.

There is a scarcity of evidence that different readability levels actually change consumers' impressions of health websites. Bates, Romina and Ahmed (2007) attempted to provide such evidence. They selected three existing websites about lung cancer, each written by trusted organizations but with differing readability levels. Participants were randomly assigned to view one of the three website and then asked to rate how credible the site was. The researchers wanted to know whether or not the website with the lowest readability score would receive the most credible ratings. Instead, the website with the middle readability score was judged to be the more trustworthy and truthful of the three. However, the researchers failed to control for other website factors, including content and design. Each of the websites provided different information about lung cancer and the preferred website was the only one to include bulleted lists, which have been shown to improve comprehension of text while raising readability scores (Redish, 2000). Having failed to isolate the difference in readability among their websites, their results are likely better explained by other factors. It is also possible that the use of bulleted lists artificially inflated the readability score of the website judged to be most credible. While the study used ecologically valid websites, the internal validity of the study suffered. The present study has the same goal as the work by Bates, Romina and Ahmed, while striving to improve on their design limitations.

The work by Bates, Romina and Ahmed (2007) hypothesized that the website with the lowest readability score would be rated as the most credible. The American Medical Association and National Institutes of Health have recommended that health information be presented

between a fourth and sixth grade level (Huang, Fang, Agarwal, Bhagat, Eloy & Langer, 2015) and the United States Department of Health and Human Services has described easy to read materials as health material written at or below a sixth-grade reading level (Patel, Sanghvi, Cherla, Baredes & Eloy, 2015). Online tools for calculating readability scores often provide tools and suggestions to 'improve' readability by rewriting portions to a lower readability level. The prevailing assumption appears to be that lower readability levels are inherently more preferred than higher levels of text complexity. Alternative theories are possible. Parents may prefer to have health information written at a level that more closely approaches the language of medical experts, and would find a higher readability level more credible. It is also possible that there is a preferred middle ground between written material that is both easily understood and that appears to match the kind of language used by experts. In the current study, the assumption put forth by other researchers and readability tools, that lower readability levels will lead to greater understanding and that greater understanding will then lead to greater credibility, will be tested. The hypotheses of this study included:

1. Lower readability scores will produce higher credibility judgements.
2. Lower readability scores will produce higher intentions to act.
3. The relationship between readability and credibility will interact with user involvement, such that in groups with lower levels of involvement this relationship will be less strong than in groups with higher involvement.
4. The relationship between readability and intentions to act will interact with user involvement, such that in groups with lower levels of involvement this relationship will be less strong than in groups with higher involvement.

Method

The present study consisted of an online survey including a mock health website that users rated in terms of credibility and intentions to act. There were three versions of the website, each written at a different readability level, with all other visual design and content factors controlled for in order to isolate the impact of readability on credibility and intention to act.

Participants

Participants were recruited for a 3 x 3 (involvement level x readability level) factorial design. Both parents from the community and introductory psychology students from the University of Manitoba participated. The parent sample was the population of interest to this study, and represents an enduring and situational involvement group (EI & SI). The students were considered low on enduring involvement, and could not also be parents. One student group was induced into a situational-involvement (SI) group, while the other group did not have either enduring or situational involvement with the website (NI).

Parents from the community were recruited mostly through online methods. Since the study concerns Internet use, online recruiting was perceived as ecologically valid. In order to be included in the present study, participants needed to be parents of school-aged children (5-18 years old) and be living in Canada. Participants from both parent and student samples were required to have fluency in English. Participants were recruited online using posts in Facebook parenting groups. Some offline recruiting methods, including posters and business cards, were also used. Posters were placed in libraries and community health centres in the city of Winnipeg. Based on an a priori power analysis, minimums of 22 participants for each of three readability level groups, for a total of 66 parents were targeted for the current study. This power analysis was conducted using G*power software, with an ANCOVA design, $\alpha = .05$, power = .80 and

effect size = .25. Each participant group exceeded the minimum target of 22 ranging from 25 to 28 participants. Parents were compensated for their time with a five-dollar Tim Horton's card mailed to them in a process separate from the data collection in order to maintain participant anonymity.

The sample of undergraduate psychology students was recruited from the University of Manitoba subject participant pool. Two groups, the situational involvement (SI) and non-SI groups, of students were recruited to match approximately the number of subjects in the parent group. Students who were also parents were not included in the study. Students received a participation credit towards their introductory psychology course in compensation for their time.

Procedure

Participants were given a link to an online survey, created using Qualtrics software. The survey consisted of three sections: a brief demographics questionnaire, a mock website containing information about a very rare genetic condition, Fragile X, and some follow-up questions based on the website presented. After completing the survey parents were able to enter their contact information in a separate survey so that they were able to receive the Tim Horton's card and keep their personal information separate from the experimental data collected.

The information collected during the demographics questionnaires phase of the survey differed slightly between the parent and student samples. Parents were first asked a few questions to confirm their eligibility in the survey, including whether or not they were parents, the number of and age of their children, and in which province of Canada they resided. Both groups declared their gender, rated their frequency of using the Internet, how much they knew about Fragile X prior to the study, as well as indicating that they were fluent in English in order to proceed with the survey. Parents were then asked to state their level of education since

research has shown that age, gender, and education level have an influence on credibility judgements (Ferebee, 2007). Participants in the parent and student SI group then were presented with the following scenario:

"Please imagine yourself in the following scenario: You are the parent of a 7 year old child. Your child has been having troubles with school and exhibiting some odd behaviour over the past few months. After a check-up, your child's paediatrician suggests testing for Fragile X syndrome. This comes as a surprise to you, as you have never heard of this disorder before. Upon returning home from the doctor's office you search online for information about Fragile X syndrome. The next screen will show you a webpage about this disorder. Please take a few minutes to study it. Afterward you will be asked a series of questions about the information presented."

Students in the NI group were simply informed that they would view a website.

Three versions of a website, created for this study, were presented. Each version contained basic information about symptoms, causes, co-morbidities, and treatment options for Fragile X and the same layout, visual design, and basic information. However, the text for each version was written at different readability levels, specifically grades 8, 11 and 14. Participants were randomly assigned to one of the three readability level groups.

After taking time to read through the presented webpage, participants were asked to rate the credibility of the website and how likely they would have been to follow the advice given. Participants were then tested for their understanding of the information contained on the presented website, as a measure to ensure that participants had properly read the information presented to them. Participants were also asked at what reading level they believed the information was written. This reading level estimate was collected to explore how well users can

intuitively estimate readability, and how well these intuitions correlate with formal measures of readability.

Materials

Website.

The text content for the websites created for this study was adapted from actual online websites about Fragile X syndrome, including Wikipedia.org, fragilex.org, fragilexcanada.ca, fragile-x.ca, and ghr.nlm.nih.gov. Another website, kidshealth.org, was also used as a resource. Although kidshealth.org does not specifically contain information about Fragile X, in a previous study it was found to successfully explain health information at a low level of readability (Cormack, 2013). Fragile X was chosen because of its rarity and clinical significance in children. Fragile X is the most common inherited form of intellectual disability but is rare enough that most participants were assumed to not have enough prior knowledge of the disorder to have pre-conceived notions to interact with their judgements of the presented website's content.

The website contained three pages. The first page had an explanation of Fragile X and included information on prevalence, diagnosis and heritability. The second page contained symptoms of Fragile X and common comorbidities. The final page had information about common interventions and medications used to aid people with Fragile X.

The text of the websites was reviewed and edited by three graduate psychology students. Students evaluated the information content of the three websites to ensure that it was equivalent across reading levels and that the text written at each reading level appears to differ in language difficulty. Comments and suggestions from these students were amalgamated and discussed. These student evaluators had the chance to review the websites again after they had been edited.

The presented websites were designed using the web hosting service Weebly.com. The visual layout of the website was designed to appear professional, similar to other health website such as mayoclinic.org. A sample image of how the websites appeared is presented in Appendix A. The text of each of the three versions of the website is provided in Appendices B, C, and D below.

Readability.

There are several formulae to measure readability, each of which uses an equation based on the length of syllables, words, sentences, and/or paragraphs within a passage of text. Among the most popular of these are the Flesch Kincaid Grade Level (F-K), the Flesch Reading Ease (FRE), the Gunning Frequency of Gobbledygook Index (FOG), and the Simple Measure of Gobbledygook (SMOG). The formulae for each can be found in Table 1. While measuring the readability of online health information, McInnes and Haglund (2011) found that each of these three measures were highly correlated, $r > .93$ for all comparisons, and used an average of the three scores to compute a single reading grade (RG). The same method was used in the present study to determine readability during the creation of the website content. Information about the readability of the presented websites is contained in Table 2. The readability results show that the websites were written at roughly an equal increment of difficulty, with about 3 grade levels difference between the low and medium websites, as well as between the medium and high websites. Calculations of readability levels were performed using an online tool, available at http://www.online-utility.org.

All previously mentioned readability formulae are based on counting various aspects of a text passage, such as the number words, syllables, and sentences. As previously cited above, these counting measures have several criticisms (Redish, 2002; Schriver, 2002). As a

comparison, an alternative method for measuring text complexity was also calculated on the website text. This technique, called latent semantic analysis (LSA) uses computer learning to calculate the similarity of words in a text passage, rather than counting text characteristics. In LSA, adjoining words in a passage of text are compared using high-dimensional vectors to represent each word's semantic meaning (Foltz, Kitsch, & Landauer, 1998). Words with similar meanings are located closer to each other in this vector space than non-similar words. LSA calculates the distance between pairs of words in this semantic hyper-space to determine their similarity. LSA has no predetermined assumptions about the meaning of individual words but rather uses computer learning techniques to inform its measures of similarity. When using LSA, the model requires being given a large corpus of English-language text to read and study, in order to learn how individual words are related together. LSA is a powerful technique that has been demonstrated to be able to pass English fluency exams such as the TOEFL (Günther, Dudschig, & Kaup, 2015).

LSA can be used to produce a measurement of the coherence of a passage of text by calculating the semantic distance between adjacent sentences, or how well each sentence fits and flows together. "In a coherent text, the understanding of a new sentence is facilitated because most concepts used in this sentence are already part of the preceding discourse" (Günther, Dudschig, & Kaup, p. 937). Since coherence is a measure of the cosine similarity of word and sentence vectors, it can produce a value between -1 and 1, where 0 would signify that there is no similarity of meaning and therefore no coherence to the passage and 1 would signify that each sentence has the same meaning. Negative coherence values are mathematically possible, but in a practical sense are nearly impossible to produce. Coherence is intended to be a measure of text comprehension (Günther, Dudschig, & Kaup), similar to readability measures. Coherence

measures were produced for each of the three versions of the website, using the LSAfun coding package described in Günther, Dudschig, and Kaup (2015). The model's vector space, the large corpus of text used to teach the model about human language, was the LSA 100k space located at http://www.lingexp.uni-tuebingen.de/z2/LSAspaces/, containing 300 dimensions and almost 5.4 million documents. This space was selected due its extremely high volume, providing the model with a large and varied sample of written material to learn from.

The low, medium, and high readability websites produced coherence measures of 0.7237, 0.6861, and 0.6934, respectively. These coherence results suggest that the coherence of each of the websites is relatively similar, but the ordering of sites differs from the values produced by readability measures. This reinforces the fact that coherence and readability are measuring distinct textual characteristics. The low readability website produced the highest coherence measure. One possible explanation for this is that the low readability website included a larger number of sentences. Coherence measures are based on how similar in meaning each adjacent sentence is to one another. In order to present the information in a way that was easier to read for the low readability website, longer sentences that appeared in the higher readability versions were broken down into smaller component sentences. These smaller sentences would likely be highly similar, since they could also be combined to produce a singular sentence at a higher readability level. However, this possibility does not explain the fact that the medium readability website produced a lower coherence score than the high readability website. The medium readability website is also made up of a larger number of smaller component sentences compared to the high readability website.

Based on the information about the semantic nature of the text provided by the coherence scores of the websites, we would expect the low readability website to be the easiest to read from

one sentence to the next, followed by the high readability website and then finally the medium readability website.

Prior knowledge of Fragile X syndrome.

After viewing the websites for this study, participants were asked to declare their prior knowledge of Fragile X syndrome by selecting which of the following statements best completes the sentence: Before viewing this website…

I had never heard of Fragile X syndrome before.

The term Fragile X sounded vaguely familiar.

I had heard of Fragile X, but wasn't familiar with it.

I had learned a little about Fragile X.

I would have considered myself very familiar with Fragile X syndrome.

Credibility.

For the purposes of this study, credibility was measured using a five-item Likert scale grounded in Hilligoss's theory of online credibility (Hilligoss & Rieh, 2008) based on a content analysis of Internet users' search history journals. This scale was used by Lederman, Fan, Smith and Chang (2014) and showed excellent internal consistency with factor loadings on the overall score all above .88. The five items are:

I think the message is credible.

I think the message is believable.

I think the message is trustworthy.

I think the message is truthful.

I think the message is reliable.

Participants rated each item on a seven point scale from 1=Strongly Disagree to 7=Strongly Agree. After the fifth item, participants were presented with an open-ended "Why or why not?" question. This provided some qualitative information about what factors participants considered in their credibility ratings.

Intention to act.

A participant's intention to act on the recommendations of the website was measured by a three item Likert scale, based on a scale used by Hu & Shyam Sundar (2010). Their scale showed excellent internal consistency, with a measure of Cronbach's $\alpha = .87$. Cronbach's α is a measure of the correlation among test items (Cronbach, 1951). As with other measures of internal consistency, a score above .8 suggests good consistency among test items. The three items include:

I would act on the recommendations given in this website.

I would recommend this website to another person.

I would forward this website to my online acquaintances.

Participants rated the items on a seven point scale from 1 = Extremely Unlikely to 7 = Extremely Likely.

Understanding.

In order to gauge whether or not the participants had complied with the instruction to read the content of the website, as well as to check for understanding, participants were given a series of questions about content presented. There were two questions from each of the three website pages referring to information. The questions were designed for this study. They are:

1. Where does Fragile X get its name?

2. Why is Fragile X more common in males?

3. Name three common signs or symptoms of Fragile X.

4. What is another name for Strabismus?

5. Name one benefit from medications often used for Fragile X.

6. What is something an Occupational Therapist can do to help a child with Fragile X?

Results

Participant Characteristics

The online survey for parents was accessed over 200 times. However, only 77 parents (66 female) actually completed the online survey. The median birth year of parent participants was 1980, with a range from 1959 to 1991. Parents participated from seven different provinces across Canada, with the majority residing in Manitoba and Alberta, $n = 25$ and $n = 34$ respectively. The parent population from this study was relatively well educated, with about two-thirds of participants having completed some kind of post-secondary degree or certificate, $n = 51$.

A total of 161 students (128 female) from introductory psychology courses participated in the online survey. The median and mode birth year of students was 1997, with 57.7% of student participants born that year. Almost all participants (97%) indicated that they used the Internet at least daily. Only 6 students reported that they had some familiarity with Fragile X syndrome (i.e., they endorsed one of the last two options for prior knowledge) before participating in the survey. Table 3 displays parent and student characteristics in greater detail.

The reported Internet usage of parents was similar to the student participants, with 96% of parent participants, $n = 68$, reporting using the Internet at least daily. However, the amount of time spent each day using the Internet among students and parents differed. Among students, 60% stated that they used the Internet for over 3 hours a day, while 55% of parents stated that

they used the Internet for time between 1 and 3 hours each day. The difference in Internet usage among students and parents was statistically significant, $\chi^2(3) = 42.34$, $p < .001$ (Internet usage responses of 'never' or 'monthly' were not included in this comparison because these options were not chosen in at least one of the participant groups, which violates the assumptions of the chi-square test). Both groups reported using the Internet on daily basis, but students reported using it for greater periods each day.

Almost a quarter of parent participants, $n = 17$, reported having some prior knowledge of Fragile X syndrome, based on selecting either the fourth or fifth item options. This also differed from students' responses of prior Fragile X knowledge. Parents reported having greater prior knowledge of Fragile X syndrome, $\chi^2(4) = 98.73$, $p < .001$.

Parent Credibility ratings

Overall, parents rated the credibility of the websites only slightly above the neutral score of 4, with average responses ranging from 4.2 to 5.2. Parent credibility ratings are displayed in Table 4, and presented graphically in Figure 2. A One-Way ANCOVA was conducted for each of the five credibility dependent variables (credible, believable, truthful, reliable, and trustworthy), with readability as the independent variable and gender, age, education level, Internet use, and prior knowledge of Fragile X as covariates. There were no significant differences between the three readability groups. The website readability level was not a significant factor in parent's credibility ratings, with readability failing to reach significance for each analysis, $p > .05$. ANCOVA statistics for each credibility measure is contained in Table 5. Parents' prior knowledge of Fragile X was a significant, $p < .05$, predictor of 4 of the 5 credibility measures. Parents with higher reported prior knowledge of Fragile X syndrome rated the website as more credible, believable, truthful, and trustworthy, as displayed in Figure 3.

Prior knowledge only failed to reach significance for parents' ratings of the website's reliability. Effect sizes for reports of prior knowledge ranged from $\eta^2 = .070$ to $\eta^2 = .084$.

A series of post-hoc step-wise linear regressions were conducted to confirm that prior knowledge was the only significant factor in predicting parents' ratings of each credibility measure. Each linear regression included six covariates (readability, gender, age, education, Internet use, and prior knowledge), but in each model only prior knowledge was retained as a significant coefficient, $p < .05$. The strength of these models were relatively small, with measures of model variance ranging from $R^2 = .067$ to $R^2 = .127$. Further statistics for these linear regressions can be found in Table 6.

Upon inspection of parent's credibility results, as demonstrated in Figure 2, it became apparent that for each credibility measure, the medium readability version of the website was rated highest by parents. A series of post-hoc contrasts was performed to determine if the quadratic pattern that appeared in the data was statistically significant for each credibility measure. The analysis found that the pattern was not statistically significant for any of the five credibility measures. These statistics are presented in Table 7.

Parent ratings of Intention to Act

The averages of parent's ratings of their intentions to act approached the neutral response of 4, with average scores ranging from 3.4 to 4.5. A One-Way ANCOVA was conducted for each of the three intention to act measures (intention to act on information, intention to recommend the website to others, and intention to forward the website on to online acquaintances), with readability as the independent variable and gender, age, education level, Internet use, and prior knowledge of Fragile X as covariates. Readability again failed to reach significance for each analysis, $p > .05$.There were no significant differences across groups in

terms of intentions to act on the websites' content, including acting on the recommendations, recommending the site to friends or family, and forwarding to site to other online contacts or groups. None of the characteristics of the website, nor any characteristics of the parents had any significant effect on intentions to act. Parent ratings are displayed in Table 8 and Figure 4.

All analyses were conducted using JASP (JASP Team, 2016), a free statistics software program available at http://jasp-stats.org. These were then confirmed using SPSS.

Student Credibility ratings

Mean student ratings of the credibility of the mock websites fell into a similar range as parent ratings, slightly above neutral from 4.0 to 5.2. Student credibility ratings are presented in Table 9 and Figures 5 through 9. A Two-Way 2x3 (Involvement x Readability) ANCOVA was performed for each of the five credibility dependent variables (credible, believable, truthful, reliable, and trustworthy), with gender, age, Internet use, and prior knowledge of Fragile X as covariates. Neither the readability level of the website, nor the level of involvement induced before reading had any significant impact on students' credibility ratings of the website, $p > .05$. Students' prior knowledge of Fragile X syndrome produced a value of $p = .048$ for the ANCOVA analysis of the dependent variable credible. However, as the ANCOVA assumption of equal variances was violated, per Levene's test $p = .029$, it was decided that a value of $p < .01$ would be required to be considered statistically significant. There were no other significant results among student credibility ratings, $p < .05$

In similar manner as parents' credibility results, it appeared that students rated the medium readability version of the website highest in each credibility measure, as demonstrated in Figures 5 through 9. A series of post-hoc contrast tests of the quadratic pattern were conducted for each credibility measure of students' ratings. Only student ratings of how

believable the website was achieved significance, $p = .046$. All other credibility measures failed to produce a reliable quadratic pattern. Table 7 contains the statistics for these analyses.

Student ratings of Intention to Act

Student ratings were once again similar to parent ratings when considering their intentions to act on the website. Mean student ratings ranged from 3.4 to 4.4. Student ratings are presented in Table 10 and Figures 10 through 12. A series of Two-Way 2x3 (Involvement x Readability) ANCOVA analyses was performed for the three intention to act dependent variables (intention to act on information, intention to recommend the website to others, and intention to forward the website on to online acquaintances), with readability as the independent variable and gender, age, Internet use, and prior knowledge of Fragile X as covariates. There was an interaction between website readability level and involvement level induced before reading the website on student's ratings for intention to act on the recommendations presented on the website. As shown in Figure 10, students with no involvement were more likely to endorse acting on recommendations from the website when written at a lower level of readability, while students with situational involvement were more likely to endorse acting on recommendations when the website was written at a higher level of complexity. This interaction was statistically significant, $F(2, 153) = 3.428$ $p = .035$, with a small effect size, $\eta^2 = .041$ and presented in Table 11. There were no significant differences among student's intentions to recommend the website to others, or to forward the website to other online acquaintances.

Readability Measures

As previously reported in the Materials section of this study, the readability of the mock website was measured using an online tool found at http://online-utility.org to ensure that the three versions of the website were written at different readability levels. At the time of analysis,

these readability scores were verified by comparing to other online readability tools, as well as Microsoft Word which calculates the Flesch-Kincaid reading level. A Google search was conducted on October 6, 2016 with the search term "readability calculator". The first page of results provided 6 online readability measurements, as well as one duplicate result, a Wikipedia entry, a Microsoft support page, and a readability measurement (Juicy Studio) that would only calculate the readability of websites, rather than text entries. Since the mock website is not actually a published site, it was unable to be analyzed by Juicy Studio's online tool.

The three versions of the mock website were scored by each of the readability calculators on multiple reading measures. The four readability measures that have been mentioned previously in this study, and displayed in Table 1, were calculated for each of the version of the mock website. These readability measures include the F-K, the FRE, the FOG, and the SMOG. Tables 12 through 14 present the readability scores provided by each of the selected readability calculators for each readability measure.

It was found that there was some level of disagreement between each of the available readability measurement tools. Even different versions of Microsoft Word provided differing readability levels for the mock website text.

These online tools also differed in their calculations of text statistics, such as the number of characters, words, syllables, and sentences contained in the text. Readability scores are calculated based on these text statistics (see Table 1 for readability formulas). Differences in reported text statistics are also included in Tables 12 through 14.

A series of intraclass correlations were calculated for each readability measure to determine the level of agreement among the readability calculators. The intraclass correlation provides a measure of the inter-rater reliability of these instruments to calculate the readability of

text (Howell, 2010). Multiple models of intraclass correlations exist. For this study, it was assumed that the readability calculators (judges) found online were considered a random sample of all available calculators, and each calculator rated the same webpage (subject) only once. Intraclass correlations provide a score between 0 and 1, with 1 meaning perfect agreement among judges and 0 signifying a total lack of agreement. A generally accepted interpretation of intraclass correlation coefficients states that "values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability, and values greater than 0.90 indicate excellent reliability" (Koo & Li, 2015, p. 158).

Each of the readability measures other than the SMOG provided excellent levels of inter-rater reliability. For three of the four readability measures, including the F-K, $ICC(2,1) = 0.99$, the FRE, $ICC(2,1) = 0.99$, and the FOG, $ICC(2,1) = 0.99$, the readability calculators found online provided an excellent level of agreement. The readability calculators provided only a moderate level of agreement for the SMOG, $ICC(2,1) = 0.75$. While there was some minor disagreement among readability calculators, for the most part these provided an excellent level of inter-rater reliability.

Comments

After participants completed the final credibility rating (trustworthy) they were given an opportunity to complete an open-ended question to explain their credibility ratings. A preliminary qualitative analysis was performed on these comments, and some common factors were mentioned. These common factors are presented only as anecdotal evidence.

Readability.

While some participants commented that a low readability website written with plain language made things easier to understand, others thought the low readability made the site appear unprofessional.

It explained things pretty clearly and simply for anyone to understand, and gave answers to many questions. (Student, Low readability group)

It explains the syndrome really well, along with how to deal with it. It gives the science behind it in a way that is easy to understand. (Parent, Low readability group)

While the text was written simply so that most everyone could understand, it was too casual for me to take seriously. (Parent, Low readability group)

It is partly trustworthy because there is some information that sounds very convincing. On the other hand it is not, because it is not very academically written. (Student, Low readability group)

Other participants appreciated the high readability website's complexity of language.

The text certainly sounds like it was written by (or with) medical professionals. (Parent, High readability group)

Prior knowledge.

Many of the comments made about participants' prior knowledge match the statistical effect that was reported earlier. Some participants who were familiar with Fragile X syndrome found information that matched their previous knowledge. Some participants who had not heard of the disorder were sceptical of the disorder and the information contained in the website.

I know a child with this syndrome and what I read seems similar to how that child behaved. (Parent, Low readability group)

I think this message is trustworthy because I have heard of this disorder before and it sounds to be similar to what I've learned before. I also trust the message because it has lots of information. (Student, High readability group)

Although it sounded scientific and used known genetic/DNA information, it seemed like a made up illness as I had never ever heard of it before. (Parent, Low readability group)

It said that it was one of the most common genetic diseases but I've never heard of it before and I feel like it would be more talked about. (Student, Low readability group)

Appearance.

Consistent with Prominence Interpretation Theory (Fogg, 2003), many participants commented on the website's professional appearance as influencing its credibility.

I trust it because it looks like a trustworthy health website like any other health website. (Student, Low readability group)

The website looked like it was created by a clinic. (Parent, High readability group)

Because they use medical language and the website looks boring enough to be professional. (Student, Low readability group)

General distrust of the Internet.

Several participants commented that they find it difficult to trust information presented on the Internet, especially when only considering one website in isolation.

The website itself appears credible, but I am sceptical of any one website, I would have to verify by checking several other sites and talking to people who may know more. (Parent, High readability group)

Because you can't trust anything the Internet says whether or not the website looks professional or not. (Student, Low readability group)

Anything we read on the Internet could be a lie. (Student, Medium readability group)

Lack of citations.

Many participants criticized the website, correctly, for lacking any sources or citations for the information presented. Some participants noted the lack of professional accreditation as well.

I think to be more credible you would need to site the information. Where are you getting this data? Is it just the same as the info on Wikipedia? I want verification that this is the same information that my Doctor has studied or is referencing. (Parent, Medium readability group)

While everything sounded very scientific and believable there were no sources cited to back up any of the information. (Student, High readability group)

There were no studies or information verifying the information and there was no "official" or government tags on the website indicating that it had been vetted by an authority on the subject. (Student, Low readability group)

Discussion

The readability level of the mock websites did not have any reliable effect on either the credibility ratings of students and parents, or the intention to act ratings of parents. Although situational involvement was induced in students in a similar manner as in Ferebee's (2007) experiment, and parents were considered as having enduring involvement based on the hypothetical scenario involving a child with a diagnosis these results fail to replicate the results found by Sillence, Briggs, Fishwick and Harris (2007), and Ferebee (2007). In both of their studies, it was found that text features had larger effects on credibility when users were engaged with the website.

Although the experiment did not include a full 2 x 2 design in terms of levels of involvement to properly measure the interactions of situational and enduring involvement, it is nevertheless surprising that there were not more significant differences among user credibility ratings between students caused by the manipulation of situational involvement. Situational involvement was induced in participants in Ferebee's study (2007) by simply asking them to answer questions based on the content of the websites they were shown. In the present study situational involvement was induced by informing participants that they would answer questions after viewing the website, but in addition, participants were also given a hypothetical scenario that included caring for a child with a diagnosis of Fragile X syndrome. It may be valuable to shorten the instructions to more closely resemble those of Ferebee's experiment, but it was assumed that adding this hypothetical scenario would add to the situational involvement of users with the website and in turn increase the effect of involvement on participants' credibility judgements. The lack of effect from involvement in the present study puts some doubt into the replicability or generalizability of Ferebee's experiment.

The failure to replicate previous studies highlights a growing concern in the field of psychological research. A general failure to replicate results has been highlighted in the area of social psychology specifically (Abbott, 2013), and also psychology as a whole (Open Science Collaboration, 2015). One explanation for this general failure is the improperly placed incentive in psychological science. Studies are more likely to be published if they are novel, and even difficult to believe; whereas some argue that replicability, not novelty, should be the criteria for publishing (Open Science Collaboration, 2015). Replicability provides powerful evidence, whereas a failure to do so creates doubt. The long-established phenomenon of priming has come into question after failures to replicate basic experiments that have been held up as examples in

psychology courses and textbooks (Abbott, 2013). This problem gained the attention of Nobel laureate Daniel Kahneman (2012), who contacted several social psychologists warning them "I see a trainwreck looming" (p.1) in their field if they did not emphasize the importance of replication. He suggested create a replication ring of labs who conduct similar research, then proceed based on the number of labs who are able to obtain similar results.

The Open Science Collaboration (2015) is a collective of psychologists who recently undertook a project to select and facilitate replications of 100 psychological studies from major journals. Only 36 replications obtained significant results. The average effect size of the 100 studies plummeted from 0.403 to 0.197 in the replications. The results of this project have created heavy doubts into the published body of psychological work, and open several questions as to how the process of publishing and disseminating scientific results should continue.

The present study is unique in its target at studying the relationship between readability and credibility judgements, and its emphasis on internal validity and control of website features. However, the present study was also based on the previous work of other scientists. Their results were expected to be replicated as the foundation of this study. The lack of effect of involvement, and readability in general was unexpected.

One important difference between the present study and the previously cited work by Ferebee (2007) is the inclusion of a measure of prior knowledge of the website's topic in the present study. This proved to be an essential addition, as it was one of the few variables to reliably explain the responses of participants. In Ferebee's study, the enduring involvement group was composed of financial planners, who would naturally have high levels of prior knowledge about the content of financial websites as well as an enduring interest in the subject.

It is recommended that more rigorous instruments to measure this concept be created for future research. The measure used in the present study to quantify users' prior knowledge of Fragile X syndrome was created for the study. Participants chose one of five statements that best completes the sentence: Before viewing this website…

I had never heard of Fragile X syndrome before.

The term Fragile X sounded vaguely familiar.

I had heard of Fragile X, but wasn't familiar with it.

I had learned a little about Fragile X.

I would have considered myself very familiar with Fragile X syndrome.

The intent of the measure was to provide an informal sense of which participants were already familiar with the content of the mock website, which was based on material from actual websites about Fragile X. It was not anticipated that this measure would produce reliable effects on credibility judgements. The measure was treated as an interval measure, but it is highly unlikely that the theoretical distance between item responses is of equal interval. The difference between Fragile X sounding vaguely familiar and not being familiar with it, and the difference between learning a little and being very familiar are not of equal weight. In order to continue to investigate the effect of prior knowledge properly, a more robust instrument must be developed that better conforms to the assumptions of statistical testing.

It is vital to acknowledge the presence and potential of prior knowledge, as it relates to credibility judgements. Once a subject has prior knowledge of a subject that they later presented with, they are more likely to recognize and judge the information as favourable. This has been demonstrated in the illusion of truth: the repetition of information increases the perceived credibility of it (Ozubko, Fugelsang, 2011). Online web users who have previously seen

information about a particular topic are more likely to find the information believable than when they were first presented with it. The illusion of truth has also been matched to fluency and familiarity with information, not simply repetition (Ozubko, Fugelsang). This has potential implications not only for the quantification of prior knowledge, but also for the writing style and readability of online health websites. Health websites written in a language and readability level that users are more familiar with may be rated as more credible than websites written at other readability levels.

In general, results from the experiment showed that the study's hypotheses were not supported. These hypotheses were based on an overall trend from agencies and websites to recommend writing to a low readability grade level. The evidence provided by this study did not support this trend towards low readability material as a way of promoting credibility. Reading involves an interaction between the reader and the text. The study's hypotheses were based on the assumption that manipulating characteristics of the text would alter how readers believed and acted on the information presented to them from the mock health website. Instead, results showed that characteristics of the reader, rather than the text, influenced credibility ratings and intentions to act. Even when these reader characteristics were controlled for using statistical techniques, text characteristics failed to produce the expected effects.

Parents with greater knowledge of Fragile X syndrome prior to the study tended to rate the website as more credible than parents who had little or no knowledge of Fragile X. Both parents and students with little prior knowledge mentioned a level of scepticism for a disorder they had not heard of. Participants with familiarity of the disorder commented that they recognized the content, which was based on information from actual websites. Also, Fragile X is a relatively unknown disorder with an odd name, which may have contributed to the lack of

credibility among parents with little or no familiarity with the disorder. In a study involving a less obscure topic the effect of prior knowledge may have been lessened.

The effect of prior knowledge on credibility suggests the value of putting out more basic information about mental health disorders to the general public, perhaps through public service announcements or advertising, so that they can better recognize proper information if they later search for it online.

A notable interaction was detected in the effect of both website readability level and situational involvement on students' intentions to follow the website's recommendations. Students without the inducement of situational involvement were more likely to act on the recommendations of the website written at a low readability grade level, while students induced with the hypothetical scenario of caring for a child with the disorder were more likely to follow the recommendations of the complex readability level website. This contradicts one of the study's hypotheses, namely that users with greater involvement would be more likely to act on recommendations given by websites written at a lower readability grade level. The assumption behind this hypothesis was that users who were involved would follow recommendations that were easier to understand. Instead, the opposite was found. Students without involvement were more likely to follow easy to understand recommendations. Students who were considering the information to care for a hypothetical child were more likely to act on recommendations when the website was written at a more complex language level. Perhaps students who were searching the mock website with a purpose preferred to follow the advice when written in a more professional and scientific manner. It should be noted that students in this context did not rate the high readability websites as more credible, but they, nevertheless, indicated a greater intention to follow the advice found on the websites with more complex text. However, it should

also be noted that parents, who experienced both enduring and situational involvement, did not show this pattern.

The use of readability measures to evaluate the effective writing of websites has been highly criticised (Reddish, 2002; Schriver, 2002). The present study has provided a new criticism for at least one of these measures; there is a level of disagreement among the various readily obtained readability calculators. None of the readability calculators used in this study produced the exact same Flesch-Kincaid reading level score for the mock website text, with scores differing within a range of 1 to 1.5 reading levels. Despite this, there was an excellent level of inter-rater agreement for the Flesch Kincaid reading level of the mock website versions. Scores for SMOG reading levels differed to a greater degree, with some calculators providing a discrepancy of three or more reading levels. If the reliability of these readability calculators is in question concerning the SMOG, than their validity is surely suspect and their use is not recommended.

The failure to agree on the number of basic characteristics of English writing also creates some doubt on the ability to automatically assess written language. It was found that the readability calculators used in this study disagreed on how to count various text statistics, including the number of words, sentences, and syllables contained in the mock website text. Readability measures are counting measures by design, so their accuracy relies on a precise count of text statistics. The English language is complex, with several phonologically and morphologically exceptional words that do not conform to the general rules of English spelling. This makes the use of any automated method for counting syllables a very difficult technique that is likely to produce errors. However, the readability calculators used in the present study also provided different results for the number of words or sentences, which follow very

consistent rules. Calculating these statistics by hand for any large body of text would be tedious, time consuming, and prone to human error as well due to the large working memory resources required.

A final criticism provided about the limitations of readability measures is a commentary on their very nature as linear models. The popular readability measures highlighted in this study are each linear equations based on different text variables, such as the number of words, syllables or sentences, with differing weights of importance. However, it is unlikely that human vocabulary develops at a linear level. In a linear pattern, the difference between the vocabulary of a pair of children in grades 2 and 3 should be the same difference as between a pair of children in grades 9 and 10. However, children do not develop at a linear, constant rate throughout their life. There are, instead, periods of acceleration and plateaus that are not represented in a linear trend. This, once again, questions the validity of the nature of these measures.

Given the many documented limitations of readability measures, as well as those added by this study, it is disappointing to see the number of research articles investigating the readability of health information continue to rise. With the finite amount of time, energy, and funding available for research, particularly in the vital domain of healthcare, it is important to question whether the resources allocated to this research is being invested wisely. Perhaps these researchers could utilize other measures of the effectiveness of written health material, including more involved measures such as focus groups, or with more complex measures of text complexity, such as LSA.

The use of LSA techniques, which do not rely on counting text statistics, for future research provides several potentially fertile avenues for study. Because LSA models depend on being given a large corpus of written text in order to learn the associations of words in text, it

would be interesting to see if manipulations to this corpus could predict the preferences of different population groups. In the present Internet era, more and more specialized information sources are available. With so many streams of media available, it is possible for users to only be presented with information that already fits their beliefs and biases, so that their view is not contradicted or challenged. In a future study, coherence measures produced by LSA models based on different information sources, such as parent blogs and academic textbooks, could be compared to parent and student credibility judgements to determine if LSA models can predict different population credibility judgements.

Particularly in this current election cycle for the United States, there has been a great divide among stories presented by media outlets from each side of the political spectrum. Internet message boards have also played a role in reinforcing the false beliefs and conspiracies put forth by the right-wing media. It would be interesting to conduct an experiment that involved giving a large corpus of exclusively left or right wing media to an LSA model and then calculating the coherence of speeches from the candidates from both parties to see how consumption of various media streams may predict how credible a candidate's arguments appear.

A significant limitation of the present study is that, due to the difficulties in recruiting large numbers of parent participants in a timely and cost-effective manner, only one sample of parents was obtained. Twice as many students were recruited; one group was induced with situational involvement through the hypothetical scenario presented directly before being shown the mock website. To properly study the effects of involvement in a 2 x 2 (enduring x situational involvement) design it would require one more group of parents who were not presented with the hypothetical scenario so that they had enduring but not situational involvement.

The results of this study may not fully generalize to the Canadian population. The parents in our study were likely more educated than the general population. Only one participant had failed to complete a high school degree, and two-thirds of the parents had completed at least some form post-secondary degree or certificate. However, it is possible that the population of parents in this study do represent to a greater degree the population of parents who use the Internet on a regular basis. Parents of higher socio-economic status, including education level, are likely to have greater access to the Internet than parents of lower education level. Also, while parents participated from one coast of Canada to the other, the vast majority of participants were located in one of only two provinces, Alberta or Manitoba.

The present study included only one experiment, but highlights several areas that could be investigated in future work. Further work should be made to attempt to replicate the effect of involvement on credibility judgements of content factors, as presented by Ferebee (2007). Also, as mentioned previously, it is possible that the rare and novel nature of Fragile X syndrome contributed to users' credibility ratings. The present study could be replicated with a mock website centering on a more common disorder, such as ADHD, autism or Tourette's syndrome, to see if similar results are found.

The methodology of this experiment, which was designed to enhance internal validity compared to previous studies using actual websites, could also be replicated to more empirically investigate the effects of various website variables, whether design elements or content factors as in the present study. Creating mock websites with manipulated variables and using random assignment to groups can allow researchers to isolate the impact of specific website variables. Rather than simply contrasting the effects of design elements versus content factors as a whole, researchers could more specifically study the effect of a specific variable of design or content.

The present study did not find any reliable effects of readability on the credibility of Internet users. These results question the assumption that readability influences the believability of a website. However, some participants did comment on the readability of the website, for good and for bad. This shows that some participants were aware of and influenced by readability, even though this influence was not reliable across the group and was not in any unified direction. For some, a low level of readability was a hindrance to taking the message seriously, while other participants appreciated explanations that were easy to understand. Also, readability did impact students' intentions to follow recommendations, in conjunction with how involved they were when reading the website.

Conclusion

Reading comprehension is a dynamic process that involves factors of the reader as well as of the material presented. In this study, characteristics of the material were controlled for, including the use of consistent content and varying readability levels, and statistical methods to control for reader characteristics; specifically, gender, age, education, Internet usage, and previous familiarity with Fragile X syndrome. The context of the material and the involvement level of readers were also manipulated. The intention of the experiment was to highlight the contributions to credibility and intentions to act of one specific material characteristic: readability. Readability, in an interaction with involvement, did have an effect on the intentions of students to follow the recommendations made in the text material. When students' were involved in a hypothetical scenario to care of a child with a diagnosis of the disorder, they were more likely to follow recommendations when the text was written at a high level of complexity. In all other cases, contrary to the study's hypotheses, it was a characteristic of the readers, their prior knowledge of the website's topic, which influenced ratings of credibility. This study

highlights the importance of understanding and recognizing characteristics of the intended audience when producing online content.

References

Abbott, A. (2013, April). Disputed results a fresh blow to social psychology. *Nature, 497*(7447). Retrieved from www.nature.com.uml.idm.oclc.org

Akram, G., Boyter, A., & Thomson, A. (2010). Evaluation of antipsychotic information on UK schizophrenia websites. *The Psychiatrist, 34*(10), 422-426.

Bates, B. R., Romina, S. M., & Ahmed, R. (2007). The effect of improved readability scores on consumers' perceptions of the quality of health information on the internet. *Journal of Cancer Education, 22*(1), 25-26.

Begeny, J. C., & Greene, D. J. (2014). Can readability formulas be used to successfully gauge difficulty of reading materials? *Psychology in the Schools, 51*(2), 198-215.

Breckon, M., Jones, R., Morris, J., & Richardson, J. (2008). What do evaluation instruments tell us about the quality of complementary medicine information on the internet? *Journal of Medical Internet Research, 10*(1), 36-50.

Brigo, G. Otte, W. M., Igwe, S. C., Tezzon, F., & Nardone, R. (2015). Clearly written, easily comprehended? The readability of websites providing information on epilepsy. *Epilepsy & Behavior, 44,* 35-39.

Burneo, J. G. (2006). An evaluation of the quality of epilepsy education on the Canadian world wide web. *Epilepsy & Behavior, 8*(1), 299-302.

Canadian Literacy and Learning Network. (2012). *Literacy Statistics.* Retrieved November 12, 2012, from http:// www.literacy.ca/literacy/literacy-sub/

Cheng, C., & Dunn, M. (2015). Health literacy and the Internet: A study on the readability of Australia online health information. *Australian and New Zealand Journal of Public Health,* Published online February 1, 2015. doi:10.1111/1753-6405.12341

Cormack, D. T. (2013). *Do attention-deficit hyperactivity disorder websites make the grade? A rubric to assess online health information.* (Unpublished honour's thesis). University of Manitoba, Canada.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

Ferebee, S. (2007). An examination of the influence of involvement level of web site users on the perceived credibility of web sites. In Y. de Kort, W. IJsselsteijn, C. Midden, B. Eggen, & B. J. Fogg (Eds.), *Persuasive Technology SE – 23* (Vol. 4744, pp. 176-186). Springer Berlin Heidelberg. doi:10.1007/978-3-540-77006-0_23

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32,* 221-233.

Fogg, B. J. (2003). Prominence-interpretation theory: Explaining how people assess credibility online. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 722-723). New York, New York, USA: ACM Press.

Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., … Treinen, M. (2001). What makes web sites credible? A report on a large quantitative study. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 61-68). New York, New York, USA: ACM Press.

Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., & Tauber, E. R. (2003). How do users evaluate the credibility of web sites? A study with over 2500 participants. *Proceedings of the 2003 conference on designing for user experiences* (DUX'03). San Francisco, California, USA: ACM Press.

Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes, 25*(2-3), 285-307.

Fox, S. (2007). *E-patients with a disability or chronic disease*. Retrieved from Pew Internet Research website: http://www.pewinternet.org/2007/10/08/e-patients-with-a-disability-or-chronic-disease/

Gunning, R. (1952). The technique of clear writing. New York: McGraw-Hill.

Günther, F., Dudschig, C., & Kaup, B. (2015). LSAfun – An R package for computations based on latent semantic analysis. *Behavior Research Methods, 47*(4), 930-944.

Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing and Management, 44,* 1467-1484.

Howell, D. (2010). Intraclass correlation. In Howell, D. (Author), *Statistical methods for psychology, seventh edition* (pp. 495-498). Belmont, CA, USA: Cengage Wadsworth

Hu, Y., & Shyam Sundar, S. (2010). Effects of online health sources on credibility and behavioral intentions. *Communication Research, 37*(1), 105-132.

Huang, G., Fang, C. H., Agarwal, N., Bhagat, N., Eloy, J. A., & Langer, P. D. (2015). Assessment of online patient education materials from major ophthalmologic associations. *Journal of American Medical Association Ophthalmology,* Published online February 05, 2015. doi:10.1001/jamaophthalmol.2014.6104

Kahneman, D. (2012, September 26). A proposal to deal with questions about priming effects. [Electronic mailing list message]. Retrieved from http://www.nature.com/polopoly_fs/7.6716.1349271308!/suppinfoFile/Kahneman%20Letter.pdf

Koo, T. K., Li, M. Y. (2015). A guideline of selecting and reporting intraclass correlations coefficients for reliability research. *Journal of Chiropractic Medicine, 15,* 155-163.

McInnes, N., & Haglund, J. A. (2011). Readability of online health information: Implications for health literacy. *Informatics for Health & Social Care, 36*(4), 173-189.

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading, 22,* 639-646.

Mead, N., & Bower, P. (2000). Patient-centredness: A conceptual framework and review of the empirical literature. *Social Science & Medicine, 51*(7), 1087-1110.

Mitchell, J., & Read, J. (2011). Attention-deficit hyperactivity disorder, drug companies and the internet. *Clinical Child Psychology and Psychiatry, 17*(1), 121-139.

Murphy, R., Frost, S., Webster, P., & Schmidt, U. (2004). An evaluation of web-based information. *International Journal of Eating Disorders, 35*(2), 145-154.

Naval Technical Training Command Millington TN Research Branch. (1975). *Derivation of new readability formulas (Automated readability index, Fog count and Flesch reading ease formula) for Navy Enlisted Personnel* (Accession Number: ADA006655).

Oermann, M. H., Lowery, N. F., & Thornley, J. (2003). Evaluation of web sites on management of pain in children. *Pain Management Nursing, 4*(3), 99-105.

Olteanu, A., Peshterliev, S., Liu, X., & Aberer, K. (2013). Website credibility: Features exploration and credibility prediction. *Proceedings of the 35th European conference on Advances in Information Retrieval* (pp. 557-568). Berlin, Heidelberg: Springer-Verlag.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), doi: 10.1126/science.aac4716.

Ozubko, J. D., & Fugelsang, J. (2011). Remembering makes evidence compelling: Retrieval from memory can give rise to the illusion of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 270-276.

Patel, C. R., Sanghvi, S., Cherla, D. V., Baredes, S., & Eloy, J. A. (2015). Readability assessment of internet based patient education materials related to parathyroid surgery. *Annals of Otology, Rhinology & Laryngology*, doi: 10.1177/0003489414567938.

Plantin, L., & Daneback, K. (2009). Parenthood, information and support on the Internet: A literature review of research on parents and professionals online. *BMC Family Practice, 18*, 10–34.

Redish, J. G. (2000). Readability formulas have even more limitations than Klare discusses. *ACM Journal of Computer Documentation, 24*(3), 132-137.

Reynolds, K. A., Walker, J. R., Walsh, K., & The Mobilizing Minds Research Group. (2014). How well do websites concerning children's anxiety answer parents' questions about treatment choices? *Clinical Child Psychology and Psychiatry,* doi: 10.1177/1359104514534948.

Schriver, K. A. (2000). Readability formulas in the new millienium: What's the use? *ACM Journal of Computer Documentation, 24*(3), 138-140.

Schwarz, J., & Morris, M. R. (2011). Augmenting web pages and search results to support credibility assessment. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1245-54). New York, New York, USA: ACM Press.

Sillence, E., Briggs, P., Fishwick, L., & Harris, P. (2007). Trust and mistrust of online health sites. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 663-670). New York, New York, USA: ACM Press.

Sillence, E., Briggs, P., Harris, P., & Fishwick, L. (2006). Changes in online health usage over the last 5 years. *CHI 06 Extended Abstracts on Human Factors in Computing Systems CHI 06,* 1331-1336.

Sobota, A., & Ozakinci, G. (2015). The quality and readability of online consumer information about gynecologic cancer. *International Journal of Gynecological Cancer, 25*(3), 537-541.

Statistics Canada, Employment and Social Development Canada, and Council of Ministers of Education, Canada (2013). *Skills in Canada: First Results from the Programme for the International Assessment of Adult Competencies (PIAAC)* (Catalogue no. 89-555-X). Retrieved from http://www4.hrsdc.gc.ca/.3ndic.1t.4r@-eng.jsp?iid=31

The JASP Team. (2016). JASP (Version 7.1.12) [Computer software]. Retrieved from http://jasp-stats.org

Turner, A., Kabashit, A., Guthriet, H., Burket, R., & Turner, P. (2011). Use and value of information sources by parents of child psychiatric patients. *Health Information and Libraries Journal, 28*, 101–109.

Underhill, C., & McKeown, L. (2008). *Getting a second opinion: Health information on the Internet.* (Catalogue no. 82–003-X Health Reports, 19, 65–69). Retrieved from Statistics Canada website: http:// www.statcan.gc.ca/pub/82-003-x/2008001/article/10515-eng.htm

Vega, L., Montague, E., & DeHart, T. (2010). Trust in health websites: A review of an emerging field. *Proceedings of the ACM International Health Informatics Symposium* (pp. 700-709). New York, New York, USA: ACM Press.

Table 1

Readability measures

Name	Formula	Source
F-K	Grade = 0.39 (average words/sentence) + 11.8 (average syllables/word) – 15.59	Naval Technical Training Command Millington TN Research Branch, 1975
FRE	FRE = 206.835 – 1.015 (average words/sentence) – 84.6 (average syllables/word)	Flesch, 1948
FOG	Grade = 0.4 [(average words/sentence) + (percent of hard words)]	Gunning, 1952
SMOG	Grade = 3 + (square root of the number of polysyllable word count)	McLaughlin, 1969

Note. F-K = Flesch-Kincaid Grade Level; FRE = Flesch Reading Ease; FOG = Gunning Frequency of Gobbledygook Index; SMOG = Simple Measure of Gobbledygook.

Table 2

Readability statistics of the mock websites

Page	Low readability website	Medium readability website	High readability website
Page 1			
FOG	8.70	11.82	14.57
SMOG	9.71	12.12	14.37
F-K	7.18	10.24	12.81
RG	8.53	11.39	13.92
Page 2			
FOG	8.02	10.39	15.13
SMOG	8.95	10.46	14.11
F-K	5.98	8.57	13.36
RG	7.65	9.81	14.20
Page 3			
FOG	9.58	12.76	16.04
SMOG	9.53	11.85	14.29
F-K	6.46	9.92	13.53
RG	8.52	11.51	14.62
Total			
FOG	8.53	11.47	15.09
SMOG	9.35	11.40	14.25
F-K	6.49	9.47	13.14
RG	8.12	10.78	14.16

Note. FOG = Gunning Frequency of Gobbledygook Index; SMOG = Simple Measure of Gobbledygook; F-K = Flesch-Kincaid Grade Level; RG = Reading Grade.

Table 3

Student and parent characteristics (percentage of sample total in parentheses)

Characteristic	Students (n = 163)		Parents (n = 77)	
Province				
Alberta			25	(32.5)
British Columbia			9	(11.7)
Manitoba	163	(100)	34	(44.2)
Newfoundland			1	(1.3)
Nova Scotia			3	(3.9)
Ontario			4	(5.2)
Saskatchewan			1	(1.3)
Education				
Less than High School/GED			1	(1.3)
High School/GED			6	(7.8)
Some University/College	163	(100)	19	(24.7)
College Degree/Certificate			15	(19.5)
Bachelor's Degree			27	(35.1)
Masters' Degree			6	(7.8)
Doctoral Degree			1	(1.3)
Professional Degree (JD/MD)			2	(2.6)
Internet Use				
Never	2	(1.2)	0	
Once a Month	0		1	(1.3)
2-3 Times a Week	3	(1.8)	2	(2.6)
< 1 Hour a Day	14	(8.6)	6	(7.8)
1-3 Hours a Day	45	(27.6)	43	(55.8)
3+ Hours a Day	99	(60.7)	25	(32.5)
Prior Knowledge of Fragile X syndrome				
I had never heard of Fragile X syndrome before.	139	(85.3)	49	(63.6)
The term Fragile X sounded vaguely familiar.	4	(2.5)	3	(3.9)
I had heard of Fragile X, but wasn't familiar with it.	14	(8.6)	8	(10.4)
I had learned a little about Fragile X.	4	(2.5)	12	(15.6)
I would have considered myself very familiar with Fragile X syndrome.	2	(1.2)	5	(6.5)

Note. GED = General Education Development test; JD = Juris Doctor degree; MD = Medical Doctor degree.

Table 4

Mean parent credibility ratings

Readability	n	Means				
		Credible	Believable	Truthful	Reliable	Trustworthy
Low	26	4.69	4.96	4.73	4.23	4.23
Medium	25	4.88	5.20	4.96	4.72	4.52
High	26	4.65	4.92	4.62	4.54	4.42

Table 5

ANCOVA analyses of parents' credibility ratings

	Credible					Believable			
	SS	df	F	η^2		SS	df	F	η^2
Readability	0.546	2	0.146	0.004	Readability	0.929	2	0.303	0.008
Gender	2.395	1	1.276	0.016	Gender	2.722	1	1.773	0.022
Age	1.505	1	0.801	0.010	Age	0.215	1	0.140	0.002
Education	0.095	1	0.051	0.001	Education	0.929	1	0.605	0.008
Internet Use	0.139	1	0.074	0.001	Internet Use	3.149	1	2.051	0.026
Prior Knowledge	11.438	1	6.093*	0.079	Prior Knowledge	8.872	1	5.779*	0.072
Error	129.523	69			Error	105.920	69		
	Truthful					Reliable			
	SS	df	F	η^2		SS	df	F	η^2
Readability	0.401	2	0.125	0.003	Readability	2.915	2	0.779	0.021
Gender	4.093	1	2.555	0.032	Gender	0.236	1	0.126	0.002
Age	1.464	1	0.914	0.011	Age	0.568	1	0.304	0.004
Education	1.244	1	0.777	0.010	Education	1.257	1	0.672	0.009
Internet Use	0.855	1	0.533	0.007	Internet Use	0.006	1	0.003	0.000
Prior Knowledge	10.819	1	6.754*	0.084	Prior Knowledge	6.784	1	3.627	0.048
Error	110.536	69			Error	129.055	69		
	Trustworthy								
	SS	df	F	η^2					
Readability	1.836	2	0.489	0.013					
Gender	0.392	1	0.208	0.003					
Age	0.060	1	0.032	0.000					
Education	0.351	1	0.187	0.002					
Internet Use	0.320	1	0.170	0.000					
Prior Knowledge	9.908	1	5.272*	0.070					
Error	129.676	69							

Note. * $p < .05$;

Table 6

Stepwise linear regression analyses of parents' credibility ratings

		Credible				
	SS	df	F	p	R	R^2
Regression	15.03	1	8.429	0.005	0.318	0.101
Residual	133.77	75				
Coefficients	Unstandardized	Standard Error	t-value	p		
Intercept	4.115	0.264	15.608	< 0.001		
Prior Knowledge	0.317	0.109	2.903	0.005		
		Believable				
	SS	df	F	p	R	R^2
Regression	10.17	1	6.704	0.012	0.286	0.082
Residual	113.78	75				
Coefficients	Unstandardized	Standard Error	t-value	p		
Intercept	4.512	0.243	18.555	< 0.001		
Prior Knowledge	0.260	0.101	2.589	0.012		
		Truthful				
	SS	df	F	p	R	R^2
Regression	17.03	1	10.94	0.001	0.357	0.127
Residual	116.76	75				
Coefficients	Unstandardized	Standard Error	t-value	p		
Intercept	4.101	0.246	16.649	< 0.001		
Prior Knowledge	0.337	0.102	3.307	0.001		
		Reliable				
	SS	df	F	p	R	R^2
Regression	9.623	1	5.401	0.023	0.259	0.067
Residual	133.623	75				
Coefficients	Unstandardized	Standard Error	t-value	p		
Intercept	3.993	0.264	15.155	< 0.001		
Prior Knowledge	0.253	0.109	2.324	0.023		
		Trustworthy				
	SS	df	F	p	R	R^2
Regression	12.20	1	6.929	0.010	0.291	0.085
Residual	132.11	75				
Coefficients	Unstandardized	Standard Error	t-value	p		
Intercept	3.826	0.262	14.604	< 0.001		
Prior Knowledge	0.285	0.108	2.632	0.010		

Table 7.

Post-hoc polynomial (quadratic) contrasts for parent and student credibility ratings

Parent ratings				
Credibility Measure	Estimate	Standard Error	t-value	p
Credible	0.028	0.283	0.100	0.921
Believable	-0.013	0.256	-0.053	0.958
Truthful	0.005	0.262	0.018	0.985
Reliable	-0.126	0.283	-0.447	0.656
Trustworthy	0.019	0.283	0.068	0.946
Student ratings				
Credibility Measure	Estimate	Standard Error	t-value	p
Credible	-0.227	0.185	-1.228	0.221
Believable	-0.339	0.168	-2.017	0.046
Truthful	-0.210	0.172	-1.223	0.223
Reliable	-0.145	0.190	-0.766	0.445
Trustworthy	-0.252	0.176	-1.438	0.153

Table 8

Mean parent ratings of intentions to act

Readability	n	Means		
		Act	Recommend	Forward
Low	26	3.88	3.81	3.46
Medium	25	4.36	4.60	4.00
High	26	4.08	4.46	4.12

Note. Act = intention to Act on website information; Recommend = intention to Recommend the website to others; Forward = intention to Forward the website to online acquaintances.

Table 9

Mean student credibility ratings

Involvement	Readability	n	Means				
			Credible	Believable	Truthful	Reliable	Trustworthy
NI	Low	27	4.63	4.81	4.96	4.33	4.33
	Medium	28	4.93	5.14	4.79	4.46	4.43
	High	25	4.68	4.68	4.36	4.28	4.08
SI	Low	26	4.19	4.19	4.08	4.15	3.81
	Medium	28	4.75	4.93	4.64	4.54	4.5
	High	27	4.70	4.70	4.48	4.44	4.41

Note. NI = no involvement; SI = situational involvement.

Table 10

Mean student ratings of intentions to act

Involvement	Readability	n	Means		
			Act	Recommend	Forward
NI	Low	27	4.15	3.93	3.56
	Medium	28	4.21	4.04	3.75
	High	25	3.52	3.52	3.24
SI	Low	26	3.65	3.58	3.34
	Medium	28	4.29	4.32	3.79
	High	27	4.48	4.37	3.89

Note. NI = no involvement; SI = situational involvement; Act = intention to Act on website information; Recommend = intention to Recommend the website to others; Forward = intention to Forward the website to online acquaintances.

Table 11

ANCOVA analyses of students' ratings of their intentions to act on the websites' recommendations.

	Act			
	SS	df	F	η^2
Readability	3.668	2	0.986	0.012
Involvement	1.145	1	0.616	0.004
Readability x Involvement	13.058	2	3.510*	0.043
Gender	1.965	1	1.057	0.006
Age	1.086	1	0.584	0.004
Internet Use	1.607	1	0.864	0.005
Prior Knowledge	3.428	1	1.843	0.011
Error	280.868	151		

Note. * $p < .05$;

Table 12

Readability scores and text statistics for the Low Readability website

Readability Measurement Tool	Readability Formula FOG	SMOG	FRE	FK
MS Word 2003-07 (Mac)	-	-	70.9	6.4
MS Word 2010 (PC)	-	-	74.8	5.8
readability-score.com	8.5	9.3	73.9	5.7
webpagefx.com	8.6	6.6	73.1	6.0
readabilityformulas.com	8.3	6.3	73.9	5.9
online-utility.org	8.5	9.4	69.7	6.5
thewriter.com	8.6	6.6	72.7	-
perrymarshall.com	8.6	6.6	-	6.0

Readability Measurement Tool	Test Statistic Characters	Words	Sentences	Complex Words[a]	Syllables
MS Word 2003-07 (Mac)	4957	1059	86	-	-
MS Word 2010 (PC)	4957	1059	86	-	-
readability-score.com	4793	1054	95	-	1540
webpagefx.com	-	1069	90	111	-
readabilityformulas.com	4793	1069	90	101	1588
online-utility.org	4825	1069	90	-	1582.1[b]
thewriter.com	-	-	-	-	-
perrymarshall.com	-	1069	90	-	1539.4[b]

Note. FOG = Gunning Frequency of Gobbledygook Index; SMOG = Simple Measure of Gobbledygook; FRE = Flesch Reading Ease; F-K = Flesch-Kincaid Grade Level. [a] – Complex Words usually denotes polysyllable words, which are used to calculate FOG and SMOG scores. [b] – Both online-utility.org and perrymarshall.com reported the average number of syllables per word. The presented syllable total was calculated by multiplying the syllable average by the number of words.

Table 13

Readability scores and text statistics for the Medium Readability website

Medium Readability Website				
	Readability Formula			
Readability Measurement Tool	FOG	SMOG	FRE	FK
MS Word 2003-07 (Mac)	-	-	51.0	9.5
MS Word 2010 (PC)	-	-	54.4	9.0
readability-score.com	10.9	11	54.9	8.5
webpagefx.com	11.6	8.7	53.6	9.1
readabilityformulas.com	11.4	8.6	53.9	9.1
online-utility.org	11.5	11.4	50.9	9.5
thewriter.com	11.6	8.7	53.1	-
perrymarshall.com	11.6	8.7	-	9.1

	Test Statistic				
Readability Measurement Tool	Characters	Words	Sentences	Complex Words[a]	Syllables
MS Word 2003-07 (Mac)	4715	888	65	-	-
MS Word 2010 (PC)	4715	888	65	-	-
readability-score.com	4549	889	77	-	1497
webpagefx.com	-	905	68	149	-
readabilityformulas.com	4549	905	68	146	1546
online-utility.org	4590	905	68	-	1520.4[b]
thewriter.com	-	-	-	-	-
perrymarshall.com	-	905	68	-	1493.2[b]

Note. FOG = Gunning Frequency of Gobbledygook Index; SMOG = Simple Measure of Gobbledygook; FRE = Flesch Reading Ease; F-K = Flesch-Kincaid Grade Level. [a] – Complex Words usually denotes polysyllable words, which are used to calculate FOG and SMOG scores. [b] – Both online-utility.org and perrymarshall.com reported the average number of syllables per word. The presented syllable total was calculated by multiplying the syllable average by the number of words.

Table 14

Readability scores and text statistics for the High Readability website

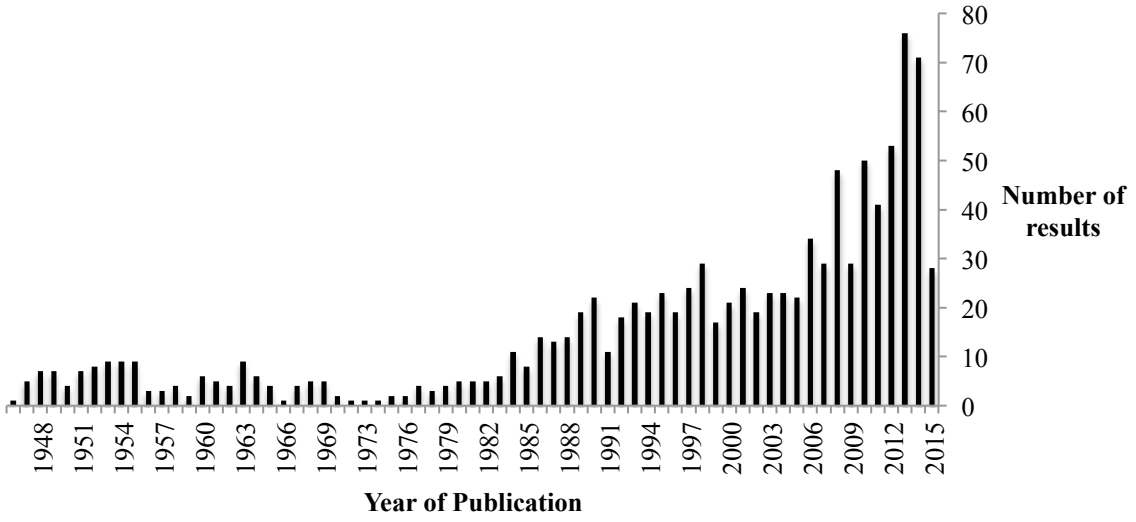
High Readability Website				
	Readability Formula			
Readability Measurement Tool	FOG	SMOG	FRE	FK
MS Word 2003-07 (Mac)	-	-	31.2	12.0
MS Word 2010 (PC)	-	-	35.1	12.4
readability-score.com	14.2	13.2	36.2	11.7
webpagefx.com	15.2	11.3	34.0	12.7
readabilityformulas.com	15.1	11.2	34.1	12.7
online-utility.org	15.1	14.2	31.0	13.1
thewriter.com	15.2	11.3	33.4	-
perrymarshall.com	15.2	11.3	-	12.7

	Test Statistic				
Readability Measurement Tool	Characters	Words	Sentences	Complex Words[a]	Syllables
MS Word 2003-07 (Mac)	4335	767	44	-	-
MS Word 2010 (PC)	4335	767	45	-	-
readability-score.com	4191	764	54	-	l441
webpagefx.com	-	780	46	174	-
readabilityformulas.com	4191	780	46	173	1457
online-utility.org	4230	780	46	-	1458.6[b]
thewriter.com	-	-	-	-	-
perrymarshall.com	-	780	46	-	1435.2[b]

Note. FOG = Gunning Frequency of Gobbledygook Index; SMOG = Simple Measure of Gobbledygook; FRE = Flesch Reading Ease; F-K = Flesch-Kincaid Grade Level. [a] – Complex Words usually denotes polysyllable words, which are used to calculate FOG and SMOG scores. [b] – Both online-utility.org and perrymarshall.com reported the average number of syllables per word. The presented syllable total was calculated by multiplying the syllable average by the number of words.

Figure 1

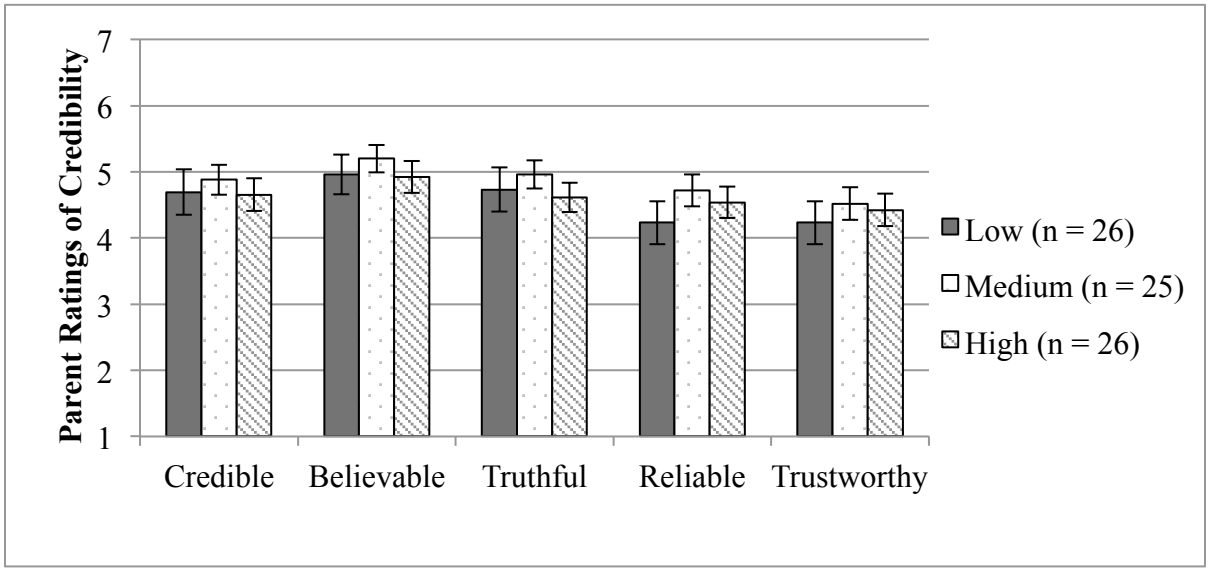
Pubmed search results for the term "flesch"



Online journal results from the Pubmed database using search term "flesch", search performed March 6, 2015. Despite criticisms of using readability formulae such as the Flesch-Kincaid Grade Level and Flesch Reading Ease, the number of studies published using these measures continues on an upward trend.

Figure 2

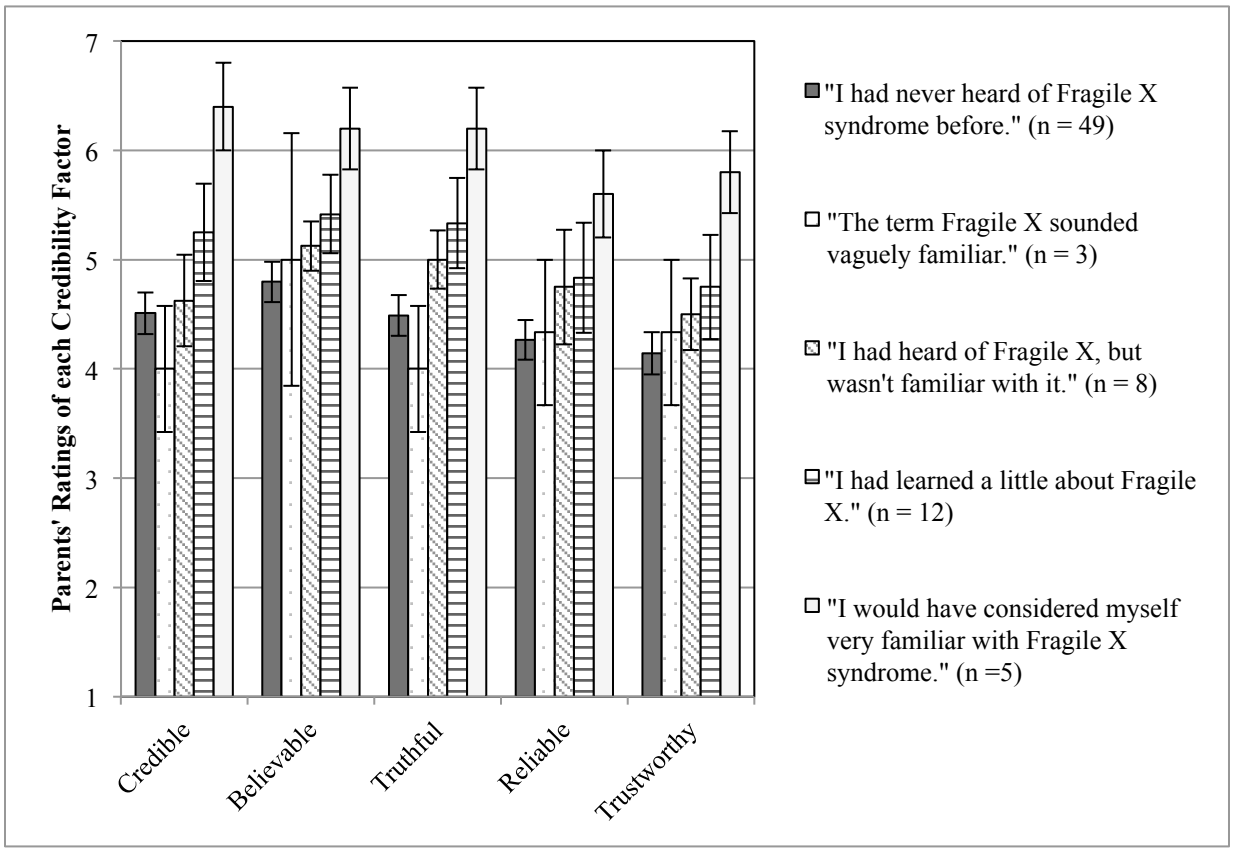
Parent credibility ratings



Error bars represent standard errors. There was no significant difference among parent responses between the three website versions. The differences in readability did not have a significant effect on parents' ratings.

Figure 3

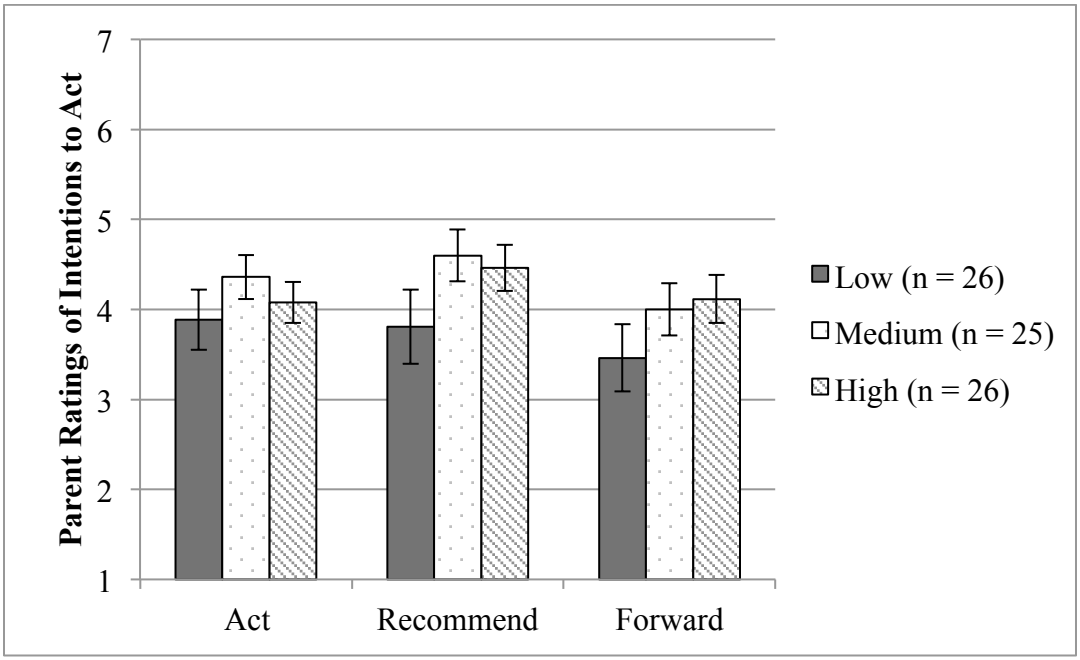
Parents' ratings of each credibility measure as a function of their prior knowledge of Fragile X syndrome



Error bars represent standard errors. Parents with greater knowledge of Fragile X syndrome prior to participating in the study rated the websites as more credible, believable, truthful, and trustworthy than parents with little or no prior knowledge of the disorder. The effect within each credibility measure, except reliable, was significant at $p < .05$, with small effect sizes ranging from $\eta^2 = .048$ to $\eta^2 = .084$. Please note the large discrepancy in group sizes. As such, these results should be interpreted with caution.

Figure 4

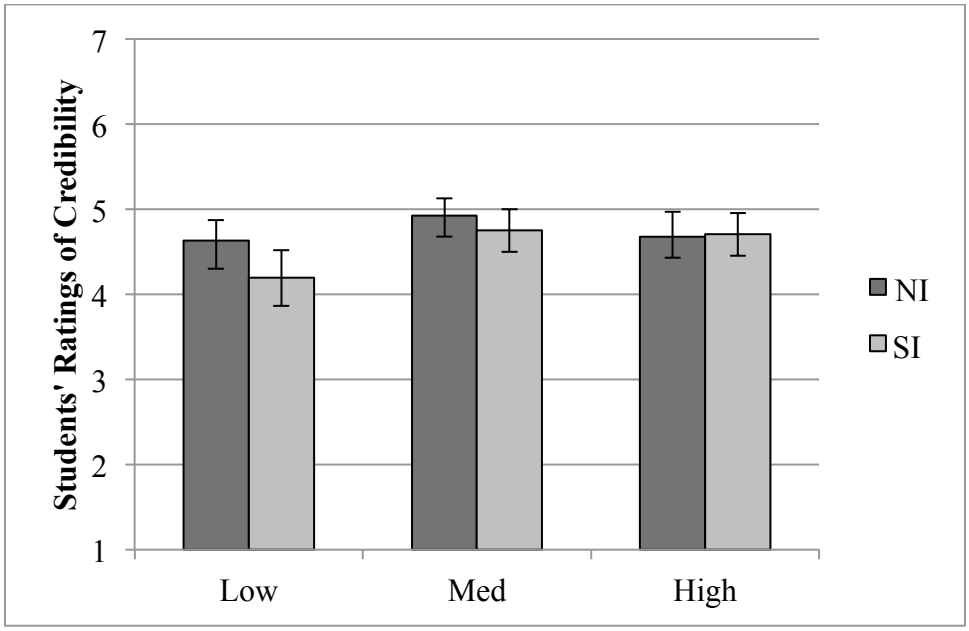
Parents' ratings of their intentions to act



Error bars represent standard error. There was no difference among readability level groups in intentions to act, with all ANCOVA's failing to reach significance, $p > .05$. Act = intention to Act on website information; Recommend = intention to Recommend the website to others; Forward = intention to Forward the website to online acquaintances.

Figure 5

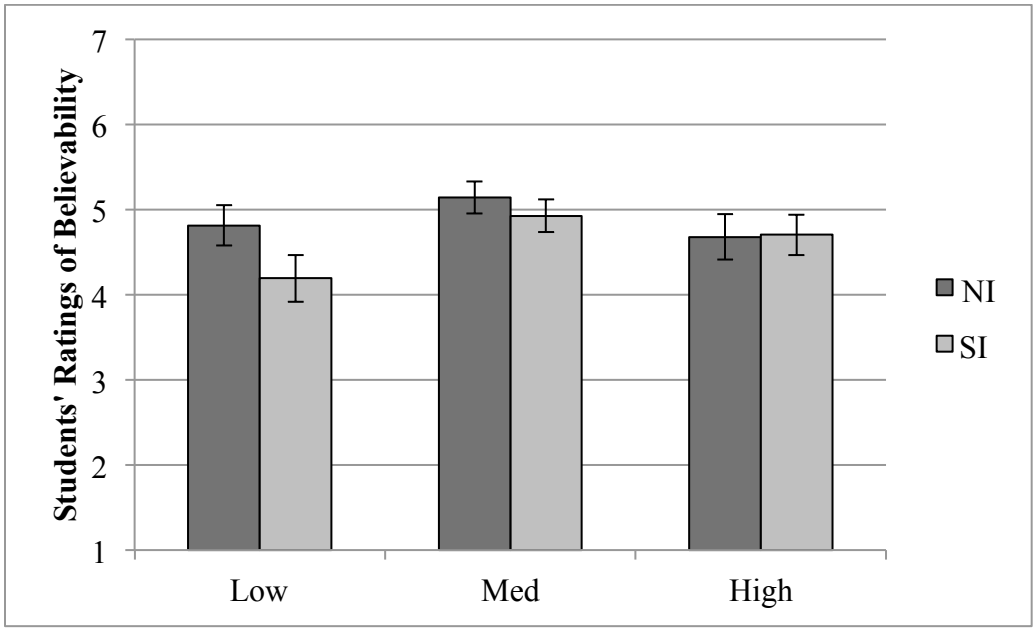
Students' ratings of the credibility of the Fragile X website



Error bars represent standard errors. As with parent participants, there was no significant difference among student responses between the three website versions. The differences in readability did not have a significant effect on students' ratings. NI = no involvement; SI = situational involvement.

Figure 6

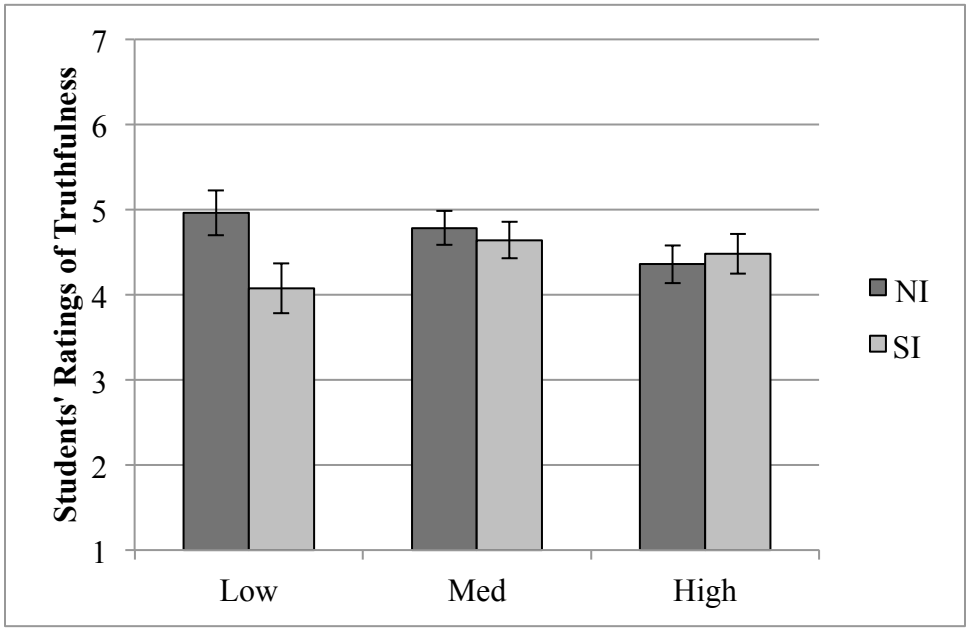
Students' ratings of the believability of the Fragile X website



Error bars represent standard errors. As with parent participants, there was no significant difference among student responses between the three website versions. The differences in readability did not have a significant effect on students' ratings. NI = no involvement; SI = situational involvement.

Figure 7

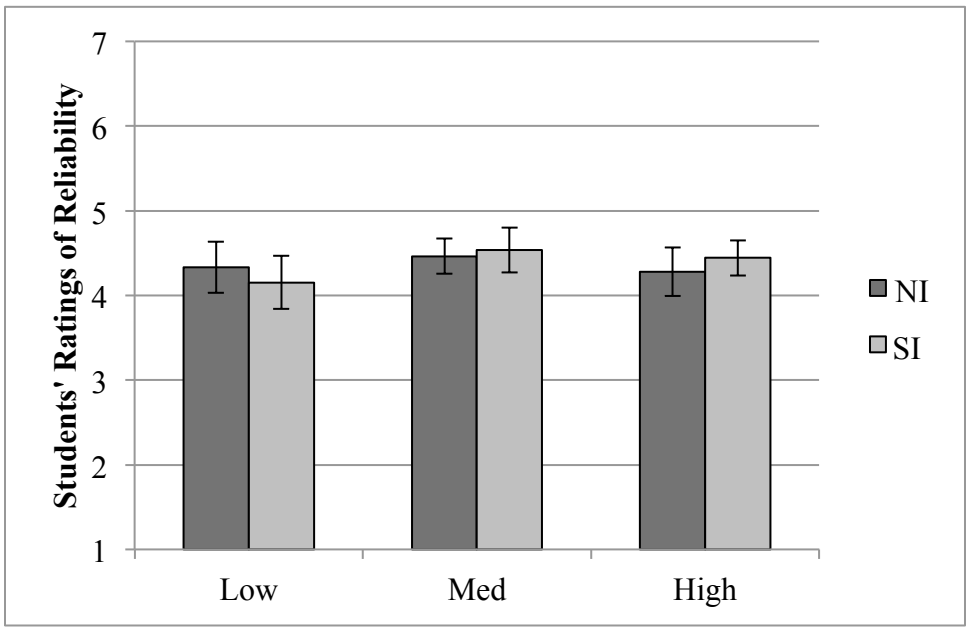
Students' ratings of the truthfulness of the Fragile X website



Error bars represent standard errors. As with parent participants, there was no significant difference among student responses between the three website versions. The differences in readability did not have a significant effect on students' ratings. NI = no involvement; SI = situational involvement.

Figure 8

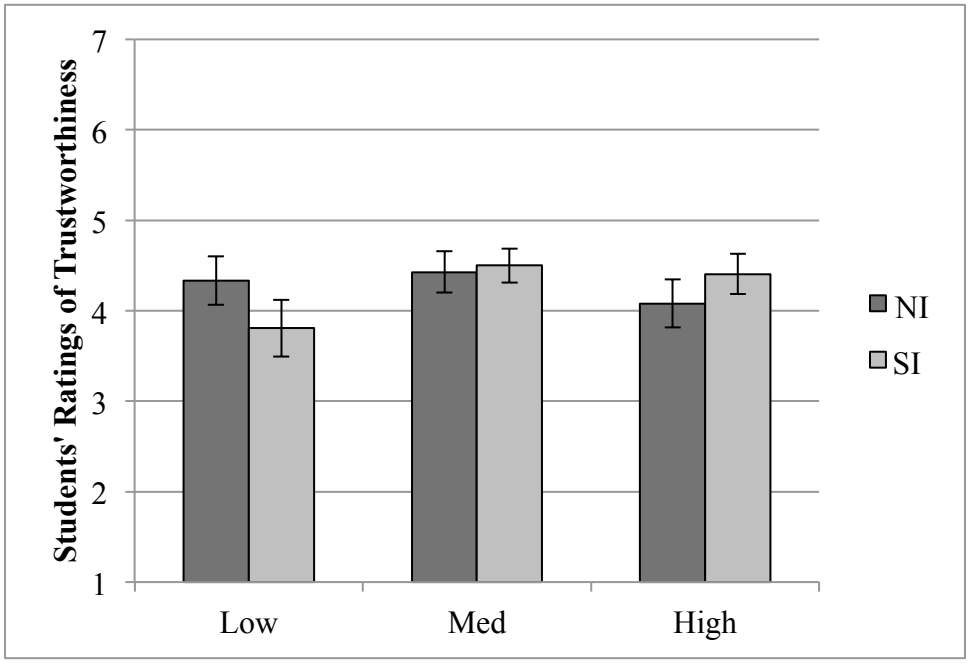
Students' ratings of the reliability of the Fragile X website



Error bars represent standard errors. As with parent participants, there was no significant difference among student responses between the three website versions. The differences in readability did not have a significant effect on students' ratings. NI = no involvement; SI = situational involvement.

Figure 9

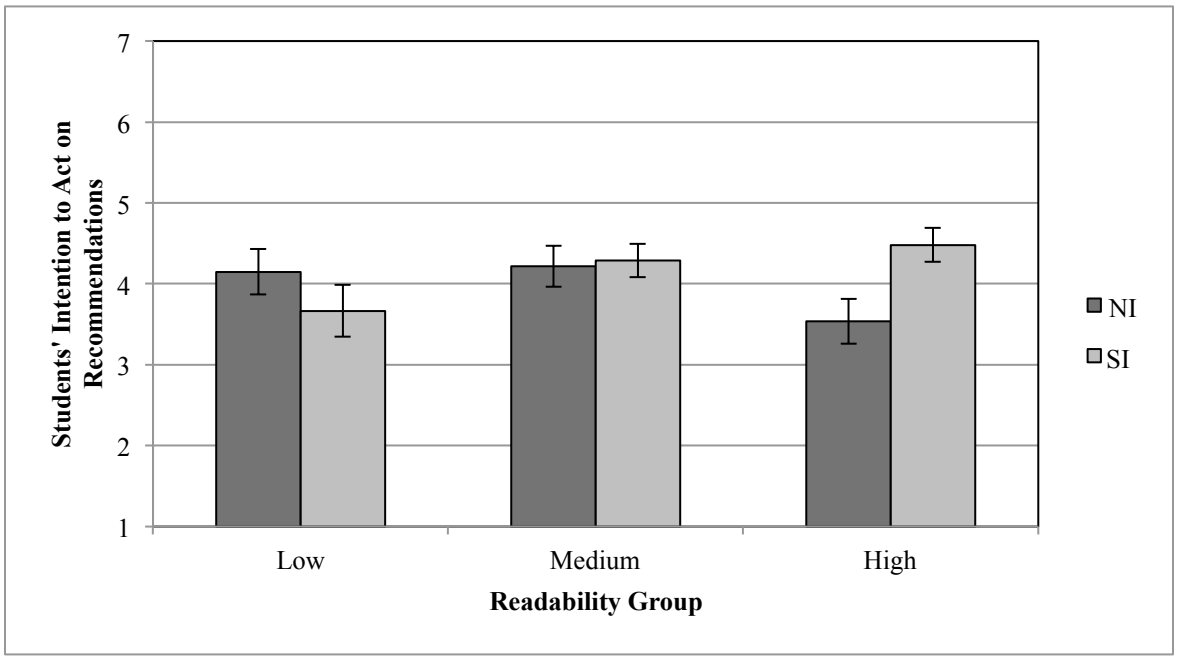
Students' ratings of the trustworthiness of the Fragile X website



Error bars represent standard errors. As with parent participants, there was no significant difference among student responses between the three website versions. The differences in readability did not have a significant effect on students' ratings. NI = no involvement; SI = situational involvement.

Figure 10

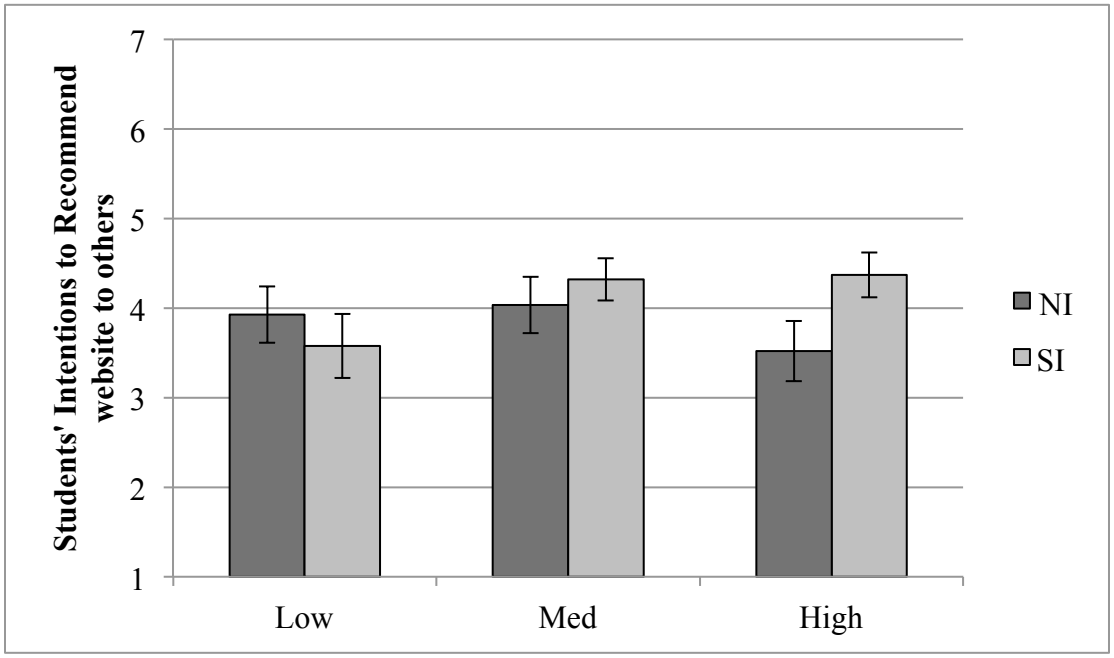
The effect of an interaction between involvement and readability level on students' intentions to act on website recommendations



Students who were induced into situational involvement were more like to follow the recommendations of websites written at a more complex level of readability. Students without any involvement with the website were more likely to follow the recommendations made on websites written at a lower level of readability. The interaction was significant with gender, age, Internet use, and prior knowledge controlled as co-variates, $p = .035$, $\eta^2 = .041$. Error bars represent standard errors. NI = no involvement; SI = situational involvement.

Figure 11

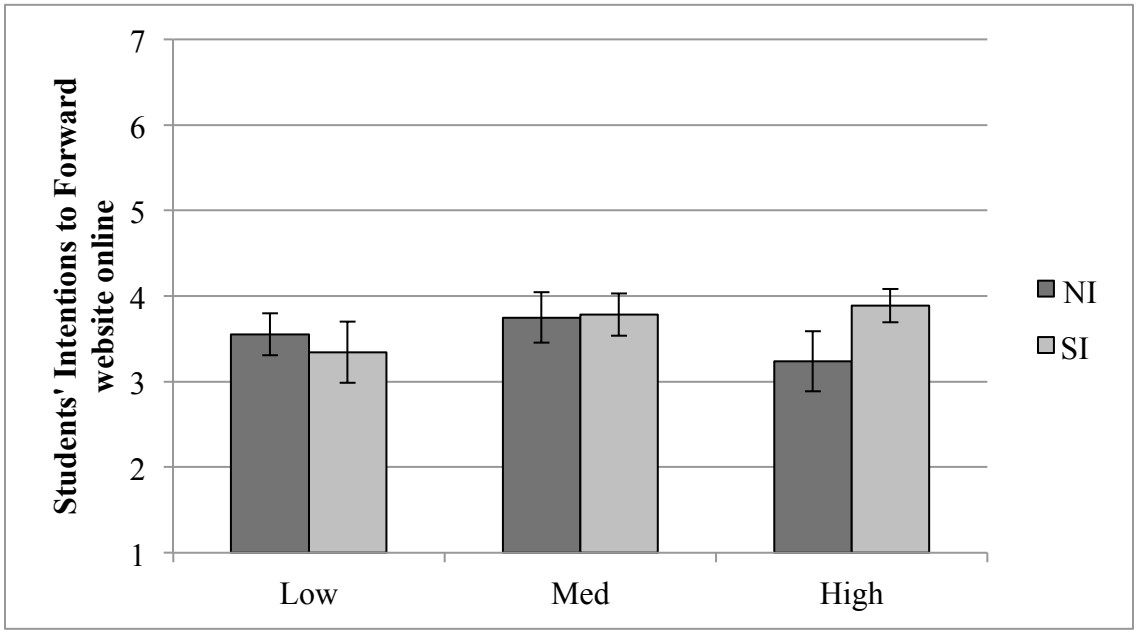
Students' ratings of their intentions to recommend the website to others



Although the results appear similar to the interaction found among students' intentions to act on website recommendations, with intentions rising with readability among students with situational involvement and declining with readability among students with no involvement (see Figure 10), the interaction was not statistically significant, $p > .05$. There were no significant differences among student responses. Error bars represent standard errors. NI = no involvement; SI = situational involvement.

Figure 12

Students' ratings of their intentions to forward the website to online acquaintances



There were no significant differences among student responses. Error bars represent standard errors. NI = no involvement; SI = situational involvement.

Appendix A

Website appearance

Search

Introduction

How Common is Fragile X?

How Does Someone Get Fragile X?

How Do I Find Out If I Have Fragile X?

What Does Fragile X Look Like?

Associated Challenges

Supports

Medicine

Fragile X syndrome is an illness that happens when someone's genes don't work right. People with Fragile X syndrome usually have troubles with thinking, behaving and learning the same way that other people do. It is the most common type of intellectual disability that can be passed on from parent to child. Fragile X is more usual in boys, and the troubles it causes are normally more serious than in girls.

How common is Fragile X?
Fragile X syndrome happens in about 1 of every 4,000 boys and 1 of every 6,000-8,000 girls. It is one of the most common diseases found in our genes. There are probably many cases of Fragile X that have not been found.

How does someone get Fragile X?
Our DNA is the blueprint of how our bodies are built. DNA is made up of 23 pairs of chromosomes that are passed on to us from our parents. Chromosomes are made up of lots of tiny genes that decide how we look. Some genes decide what colour our hair or our eyes are. One pair of chromosomes decides our gender. A mother passes on one X chromosome to her children. Children receive either an X or Y chromosome from their father. Children with two X chromosomes are girls, while boys have one X chromosome and one Y chromosome.

In people with Fragile X syndrome, at least one of their X chromosomes is mutated. The mutation involves repeated DNA patterns in the chromosome. Boys with Fragile X have more serious challenges than girls because they only have the one problem X chromosome. Girls with only one mutant X chromosome also have one healthy X chromosome, which helps to make the challenges less severe. It is very rare, but sometimes girls might have two problem X chromosomes. Their challenges would be as severe as a boy's challenges. Neither would have a healthy X chromosome to help them.

Men with Fragile X syndrome will pass the disease on to their daughters, but not to their sons. That's because their sons will only get a healthy Y chromosome, not the mutant X chromosome. Children of women with Fragile X have a 50% chance of getting the disorder, depending on whether the healthy or problem X chromosome is passed on.

How do I find out if I have Fragile X?
Fragile X is diagnosed using DNA testing. DNA testing can also detect whether a person is a carrier of the disease, and is likely to pass it on to their children. A family doctor can order a DNA test be done. This usually involves taking somebody's blood and testing it.

INFORMATION
RESOURCES
ABOUT
CONTACT

Contact us at:
questions@fragile-x.ca

Appendix B

Low Readability website

PAGE 1

Fragile X syndrome is an illness that happens when someone's genes don't work right. People with Fragile X syndrome usually have troubles with thinking, behaving and learning the same way that other people do. It is the most common type of intellectual disability that can be passed on from parent to child. Fragile X is more usual in boys, and the troubles it causes are normally more serious than in girls.

How common is Fragile X?

Fragile X syndrome happens in about 1 of every 4,000 boys and 1 of every 6,000-8,000 girls. It is one of the most common diseases found in our genes. There are probably many cases of Fragile X that have not been found.

How does someone get Fragile X?

Our DNA is the blueprint of how our bodies are built. DNA is made up of 23 pairs of chromosomes that are passed on to us from our parents. Chromosomes are made up of lots of tiny genes that decide how we look. Some genes decide what colour our hair or our eyes are. One pair of chromosomes decides our gender. A mother passes on one X chromosome to her children. Children receive either an X or Y chromosome from their father. Children with two X chromosomes are girls, while boys have one X chromosome and one Y chromosome.

In people with Fragile X syndrome, at least one of their X chromosomes is mutated. The mutation involves repeated DNA patterns in the chromosome. Boys with Fragile X have more serious challenges than girls because they only have the one problem X chromosome. Girls with only one mutant X chromosome also have one healthy X chromosome, which helps to make the

challenges less severe. It is very rare, but sometimes girls might have two problem X chromosomes. Their challenges would be as severe as a boy's challenges. Neither would have a healthy X chromosome to help them.

Men with Fragile X syndrome will pass the disease on to their daughters, but not to their sons. That's because their sons will only get a healthy Y chromosome, not the mutant X chromosome. Children of women with Fragile X have a 50% chance of getting the disorder, depending on whether the healthy or problem X chromosome is passed on.

How do I find out if I have Fragile X?

Fragile X is diagnosed using DNA testing. DNA testing can also detect whether a person is a carrier of the disease, and is likely to pass it on to their children. A family doctor can order a DNA test be done. This usually involves taking somebody's blood and testing it.

PAGE 2

What does Fragile X look like?

Fragile X looks different in boys and girls. The most common sign is an intellectual disability (ID). Kids with an ID learn and develop more slowly than other kids. This can be a mild or a serious problem. Girls with Fragile X usually only have a few problems. Boys with Fragile X usually have a much harder time. Almost all boys with Fragile X have an ID. Only half of girls with Fragile X have an ID.

People with Fragile X usually have large ears, a long face, and flat feet. Sometimes they have soft skin and joints that extend more than usual. Teenage boys with Fragile X sometime have enlarged testicles.

Sometimes kids with Fragile X act kind of like kids with Attention-Deficit Hyperactivity Disorder (ADHD) or Autism Spectrum Disorder (ASD). For example, a kid with Fragile X might look like someone with ADHD. He or she might be really hyper, act without thinking, or fidget a lot. Or, they might make strange movements or repeat sounds like someone with ASD. They might flap their hands, bang their head, or rock back and forth.

Associated Challenges

Some kids with Fragile X also have other challenges with their bodies and minds.

ADHD is very common in kids, especially boys. Kids with ADHD have trouble paying attention and staying still. ADHD shows up in most boys with Fragile X syndrome. About one out of every three girls with Fragile X also have ADHD.

ASD is another challenge that some kids face. While ASD has some special features, no two kids with ASD act exactly the same way. Kids with ASD almost always have trouble socializing with others. They might also act strangely or have odd interests. Sometimes kids with ASD get easily overwhelmed by all of the things they see and hear. About 1 in every 4 or 5 kids with Fragile X also have ASD.

Lots of people with Fragile X get really nervous in social situations (Social Anxiety). You might notice that they don't make eye contact when talking to you. It might be hard for them to recognize people they know. Kids with Fragile X have trouble making friends and holding a conversation. Sometimes they have panic attacks.

A lot of kids with Fragile X develop a lazy eye, a.k.a. strabismus, so they should see an eye doctor.

Less than a quarter of people with Fragile X get seizures. Sometimes medicine is used to control the seizures.

PAGE 3

Supports

There is no cure for Fragile X yet. But, there are still ways to help people live well with the disorder. Anyone can develop with education, therapy, and support.

Kids with Fragile X syndrome can get extra help at school. A team of specialists is available at school to help your child. They will find out what your child needs to succeed. Then they will make plans specific to meet those needs. This team might include speech therapists, occupational therapists, and school psychologists. Speech therapists help kids who have trouble saying certain sounds. Occupational therapists help protect kids from being overwhelmed by their environment. They can plan the best classroom set up for your child. The school psychologist can find out how your child learns. They can suggest ways that the teacher can work with your child's skills and difficulties.

Medicine

Doctors often prescribe medicine for people with Fragile X. Medicines can't cure Fragile X, but they can help people think and feel better. Stimulants are one kind of medicine that can help people with Fragile X. They help people pay attention, think before acting, and sit still. Anti-depressants are another type of medicine. They help people feel better and think more clearly.

Appendix C

Medium Readability website

PAGE 1

Fragile X syndrome is a genetic condition that can cause intellectual disabilities, behavioural problems, and learning difficulties. It is the most common inherited form of intellectual disability. Fragile X is most common in males, and presents more severe symptoms than females.

Prevalence

Fragile X syndrome affects, on average, one in 4,000 males and one in 6,000-8,000 females. It is estimated to be one of the most common genetic diseases. There are likely many cases of Fragile X that are undetected.

Inheritance

Each of us receives 23 chromosomes from each of our parents, 46 in total, which make up our DNA. These chromosomes become the blueprint for our developing bodies. Some genes determine our physical features. One set of chromosomes determines our biological sex. A mother passes on one X chromosome to her children. Children receive either an X or Y chromosome from their father, determining their sex. Children with two X chromosomes are female, while males have one X chromosome and one Y chromosome.

In persons with Fragile X syndrome, at least one of their X chromosomes has a mutation caused by allele pattern repetitions. This explains why the disease presents more often and more severely in males. Females with one mutated X chromosome also have one healthy X chromosome, which buffers them against the effects of the disease. Males, and females with two mutated X chromosomes, lack a healthy X chromosome.

Males with Fragile X syndrome will pass the disease on to their daughters, but not their sons, who receive a Y chromosome. Children of females with Fragile X have a 50% chance of inheriting the disorder, depending on whether the healthy or mutated X chromosome is passed on.

Diagnosis

Fragile X is diagnosed using a DNA testing technique for counting allele patterns in DNA. DNA testing can also detect whether a person is a premutation carrier, and is likely to pass on the disease to their offspring. A family doctor can request a DNA test, usually from a blood sample.

PAGE 2

Symptoms

Fragile X usually looks different in boys and girls. The most common aspect of Fragile X is an intellectual disability. This disability can come in a wide range, from a very mild challenge to very severe. Boys with Fragile X usually have a medium to high level of intellectual disability, while it's usually only mild in girls. Only half of girls with Fragile X even have an intellectual disability.

People with Fragile X usually have some unique physical traits. These often include large ears, a long face, flat feet, soft skin, hyper-extendable joints, and enlarged testes in males.

Sometimes people with Fragile X act in similar ways to people with Attention-Deficit Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD). Examples of these kinds of behaviours include being overactive and impulsive, fidgeting, or repeating words. This could also include strange movements, such as hand-flapping, head-banging, or rocking back and forth.

Associated Disorders

People with Fragile X syndrome often have other physical and mental conditions.

ADHD is a common condition among children, especially boys. People with ADHD have problems paying attention and staying still. ADHD shows up in most boys with Fragile X syndrome, as well as about 30% of girls.

People with ASD have common symptoms that can look very different in each person. These include poor social skills, strange repetitive movements, and a tendency to be overwhelmed by their senses. About 20-25% of people with Fragile X syndrome also have ASD.

Social Anxiety is very common in people with Fragile X syndrome. You might notice that they don't make proper eye contact or have trouble recognizing faces. People with Fragile X syndrome often have trouble making conversations and keeping friends. Sometimes they have panic attacks.

Strabismus, a.k.a. "lazy eye", is common in individuals with Fragile X. If your child has Fragile X, they should see an optometrist to check for lazy eye.

About 15-25% of people with Fragile X have seizures. Sometimes medicine is used to control them.

PAGE 3

Treatments

There is currently no cure for Fragile X syndrome. But, there are plenty of treatments that can help people with Fragile X live better lives. All people can improve with proper education, therapy, and support.

Most children with Fragile X syndrome can qualify for additional support at school. The school team that will plan supports for your child is composed of a variety of different

professionals. These professionals will set goals and make plans to benefit your child based on assessments of your child's abilities and needs. This team may include speech-language pathologists, occupational therapists, and school psychologists. Speech-language pathologists help children develop the physical and social skills needed for speaking. Occupational therapists can plan classroom setups that are ideal for your child to learn. They can also help reduce stimulation if your child's senses are easily overloaded. School psychologists specialize in how individual children think and learn. They can then recommend teaching strategies based on a child's specific strengths and weaknesses.

Medications

Doctors often recommend medications for people with Fragile X syndrome to help them manage their emotions and actions. The most commonly recommended drugs are called stimulants. Stimulants can help to improve concentration and make people less hyperactive and fidgety. Another common type of drug is anti-depressants. These can help people feel less anxious and reduce mood swings. They can also reduce obsessive thoughts.

Appendix D

High Readability website

PAGE 1

Fragile X syndrome is a genetic condition associated with an intellectual disability, behavioural problems, and learning difficulties. It is the most prevailing form of inherited intellectual disability. Fragile X is most common in males and presents more severe symptomatology than in females.

Prevalence

Epidemiological studies suggest that Fragile X syndrome affects 1 in 4,000 males and 1 in 6,000-8,000 females. It is estimated to be one of the most common genetic diseases. There are likely many undiagnosed cases in the population.

Inheritance

Our DNA, composed of 23 pairs of chromosomes received in combination from our parents, determines our unique genetic design. Our biological sex is defined by the combination of sex chromosomes we inherit. A mother imparts one X chromosome to her children, and a father will impart either an X or Y chromosome. Individuals with two X chromosomes are born female, while males have one of each chromosome.

Fragile X syndrome is produced by a CCG trinucleotide expansion mutation on one of the genes of an affected X sex chromosome. While the symptoms are less severe for females with one healthy X chromosome, which buffers them against the effects of the disease, males and females with two mutated X chromosomes lack a healthy X chromosome and bear the full brunt of the effects of the disorder.

Males with Fragile X syndrome will bestow the mutated X chromosome on to their daughters, but not their sons. Children of females with Fragile X have a 50% chance of inheriting the disorder, depending on whether they inherit the healthy or mutated X chromosome.

Diagnosis

Cytogenetic analysis is used to diagnose Fragile X and also to identify premutation carriers, who are likely to pass on the disease to their offspring. Physicians can request DNA testing with a blood sample.

PAGE 2

Symptoms

Fragile X symptoms tend to present differently in males and females. The most common feature of Fragile X is an intellectual disability. This can range from mild to severe; males usually range from moderate to severe. Less than one half of females with Fragile X qualify for an intellectual disability, and those who do usually have only a mild disability.

There are common physical features associated with Fragile X, including large ears, long face, flat feet, soft skin, hyper-extendable joints, and macroorchidism (enlarged testes in males).

Behaviours common in Attention-Deficit Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD), such as hyperactivity, fidgeting, impulsivity, stereotypic movements, and repetitive speech are also common in persons with Fragile X syndrome.

Comorbidities

Fragile X syndrome is associated with a number of other mental and physical disorders.

ADHD is a common disorder among children, especially boys, that is characterized by attention difficulties and hyperactive behaviour. The majority of boys with Fragile X syndrome, as well as about 30% of girls, also qualify for a diagnosis of ADHD.

ASD is a wide spectrum of conditions typified by severe social skill deficits, sensory overload, and stereotypic movements. About 20-25% of individuals with Fragile X syndrome also have a diagnosis of ASD.

Social Anxiety is one of the most common features of Fragile X syndrome. This is usually demonstrated by poor eye contact and facial recognition deficits. Individuals with Fragile X syndrome often have difficulties initiating social interaction, maintaining relationships, and can experience panic attacks.

Strabismus, also referred to as “lazy eye”, is common in individuals with Fragile X. Early screening for vision problems is recommended.

Seizures occur in about 15-25% of individuals with Fragile X. Medications are often used to control symptoms.

PAGE 3

Treatments

There is currently no cure for Fragile X syndrome. However, there are several treatments available to improve the lives of those affected by the disorder. All individuals can make progress with the aid of education, therapy, and support.

Most children with Fragile X syndrome will be eligible for educational support services. A multi-disciplinary team will assess the child’s needs, design interventions and make goals towards their improvement. The school team working to help your child will likely include professionals such as speech-language pathologists, occupational therapists, and school psychologists. Speech-language pathologists help children with delayed speech or deficits in conversational speech. Occupational therapists can help children manage their sensory needs and determine their ideal learning environments. School psychologists can determine a child’s

learning needs and make recommendations for how to manage each child's specific strengths and weaknesses.

Medications

Individuals with Fragile X syndrome are often prescribed medications to manage their behavioural and emotional symptoms. The most commonly used medications are stimulants that help control hyperactivity, impulsivity and attentional problems. Anti-depressants are often used to manage feelings of anxiety, obsessive thoughts and to stabilize people's mood.