

Should I Trust It When I Cannot See It? Credibility Assessment for Blind Web Users

Ali Abdolrahmani and Ravi Kuber

UMBC

Baltimore, MD 21250

{ aliab1, rkuber } @umbc.edu

ABSTRACT

As users become increasingly more reliant on online resources to satisfy their information needs, care is needed to ensure that these resources are credible in nature, especially if a decision is to be taken based upon the information accessed. The credibility of a web site is known to be heavily influenced by its visual appearance. However, for individuals who are blind, challenges are often faced accessing these visual cues when using assistive technologies. In this paper, we describe an observational study to examine the strategies and workarounds developed by individuals who are blind to perform credibility assessments. These are compared with those used by sighted users. Findings from the study have highlighted the relationship between accessibility and credibility. The features used to form assessments non-visually have also been identified. Insights from the study can be used to support the design of highly credible interfaces for blind screen reader users.

CCS Concepts

• Human Centered Computing → Accessibility

Keywords

Accessibility; Blind; Visually-Impaired; Web Credibility

1. INTRODUCTION

Recent advances in technology have revolutionized the ways in which information can be accessed and shared with others. As the volume of information available online increases, researchers suggest that the content available may not be subject to filtering through professional gatekeepers. This therefore raises issues surrounding the credibility or quality of content [15]. As users become increasingly more reliant on online resources to satisfy their information needs, the presence and prominence of incorrect and misleading content can have serious consequences for users [19], particularly for those who make decisions based upon the information accessed. The burden is placed on the user to assess levels of credibility [14]; however, difficulties are often faced by users during this process [15].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ASSETS '16, October 23-26, 2016, Reno, NV, USA.

© 2016 ACM. ISBN 978-1-4503-4124-0/16/10...\$15.00.

DOI: <http://dx.doi.org/10.1145/2982142.2982173>

According to Fogg et al. [7], the two key components of credibility include trustworthiness and expertise. The trustworthiness dimension captures the perceived goodness or morality of the source, while the expertise dimension captures the perceived knowledge and skill of the source. When individuals explore online content, their assessment about the credibility of the information involves both objective judgments of information quality or accuracy, as well as subjective perceptions of the source's trustworthiness, expertise, and attractiveness [9,15]. Additionally, the visual appearance of a web page plays a significant role in attracting information seekers, as well as impacting their perception of its credibility [8].

For individuals who are blind, obtaining these visual cues using screen readers can prove to be challenging. Assistive technologies, such as screen readers, are able to translate textual content from a web page into auditory or tactile format, enabling users to gain an overview of content. However, graphical information and structural layout can be difficult to perceive non-visually. As information is outputted in a linear and time-consuming fashion through a screen reader, it can be a frustrating process when attempting to traverse through content-heavy sites. Furthermore, features such as banners or menus may not vary from page-to-page, and may consequently be presented multiple times when exploring the contents of a site, thereby "overloading" the user [1]. Further details relating to the limitations associated with screen reading technologies can be found in [1,3,16].

In this paper, we describe an observational study examining the strategies and workarounds used by individuals who are blind to perform credibility assessments when using a screen reader. Findings have been compared with those of sighted users. The definition of credibility used for this study refers to the believability of some information and/or its source [11,15]. We are particularly interested in examining: (1) the non-visual cues and credibility criteria used to make judgments (termed: features); and (2) the relationship between accessibility and credibility when making assessments.

2. RELATED WORK

2.1 Credibility assessment

The process of assessing credibility has been examined by a range of researchers. Hilligoss and Rieh [10] proposed a framework of credibility assessment in which credibility is characterized across a variety of media and resources with respect to diverse information seeking goals and tasks. The researchers conducted a diary study, where participants performed various information-seeking tasks. Through a grounded theory analysis, three distinct levels of credibility judgments emerged: construct (how a person constructs, conceptualizes, or defines credibility), heuristics (which involves general rules of thumb used to make judgments of credibility applicable to a variety of situations), and interaction

(which refers to credibility judgments based on content, peripheral source cues, and peripheral information object cues). Wathen and Burkell [20] also proposed a model of credibility assessment. The researchers suggest that surface characteristics (e.g., appearance, interface design including download speed and interactivity, and organization of information) are rated once a web site is accessed. If this initial evaluation passes the user's criteria, then they will move to the next "level" of evaluation where the source and message are rated. Factors such as source expertise, competence, trustworthiness, accuracy, currency, and relevance would be taken into account. The third aspect of the process involves the interaction of presentation and content with the user's cognitive state. The researchers suggest that if these assessments meet the user's criteria for credibility, the user accepts the information as credible, and decides to evaluate the information content. If they do not, the user will likely leave the site.

Through a series of studies, Fogg et al. [5-8], have focused on determining the features of sites used for credibility assessment. In their 2003 paper [8], the researchers have described the Prominence-Interpretation Theory to model the ways in which credibility can be assessed. The theory posits that the impact an element has on perceived credibility is a product of its prominence (how likely it is to be noticed) and interpretation (what value or meaning people assign to that element) [19]. Fogg et al. [8] introduced 2684 participants to web sites, covering a range of topics including news, health, travel, e-commerce, and asked them to review and rank these in terms of credibility. Participants commented on the 'design look' of the site more often than any other feature (46.1%). Information visual design/structure (28%) and information focus (25%) were also identified more frequently by participants. Interestingly, functionality, clarity and readability were mentioned less often. In order to develop a highly credible web site, the researchers suggest investing in the visual aspects of a site. Designers should be aware that some highly prominent elements which may impact credibility, are sometimes outside of their control. Care should also be taken when making decisions about prominence, as "not everything can stand out at once". The perceived credibility of a web site is thought to hinge on these decisions [8].

Schwarz and Morris [19] examined page features which are currently difficult or impossible for end users to assess, yet provide valuable signals regarding credibility. These included examining: (1) on-page features (e.g., spelling errors, number of advertisements), (2) off-page features (e.g., awards won, sharing/hit rate of the page through social media), and (3) aggregate features (e.g., general popularity, geographical reach, expert popularity). The researchers proposed visualizations to augment search results and web pages using the most promising of these features. Findings revealed that augmented search results were found to be particularly effective at increasing the accuracy of users' credibility assessments.

The importance of using visual aesthetics to create favorable first impressions of a site has been described by Robins and Holmes [17]. The researchers presented the same content to participants in their study, using different levels of aesthetic treatment. The content with a higher aesthetic treatment was judged as having higher credibility. Kim and Moon [13] found that it is possible to manipulate the visual design factors of an interface in order to induce a target emotion, such as trustworthiness. They suggest that interfaces should be designed which create trustworthy feelings among users, which in turn will influence the decision to use the system. However, Blythe et al. [2] highlight that there may be instances where interfaces on the surface may appear

trustworthy, but may need further checks to verify their intent. Difficulties were faced by participants in their study when attempting to detect 'phish' when visual cues, such as professional-looking logos, were present on an interface.

2.2 Information seeking by sighted and blind web users

Studies have been conducted examining the ways in which information seeking habits differ between individuals who are sighted and blind/visually impaired, and the features influencing impressions of the sites accessed. Examples include the study by Craven and Brophy [4], where the researchers presented four information seeking tasks using four different electronic resources. Results confirmed that it took visually impaired participants longer to complete searching and browsing tasks, with times varying considerably depending on the design of the site. Search time was impacted when encountering pages which contained more information, or ones that contained a number of hyperlinks. Ivory et al. [12] presented web-based search tasks to ten sighted and six blind participants. Findings from their study showed that participants initially used the page's summary, title, and URL to predict search result relevance. They then considered additional features (words, ads, and quality) to decide whether or not to explore the page, regardless of their relevance predictions. Similar to [4], blind participants were found to spend more time on tasks. They spent on average twice as long as sighted participants to explore search results and three times as long to explore web pages. Sahib et al. [18] found that the average number of results viewed by sighted participants was significantly higher than visually impaired participants. Sighted participants were also found to submit significantly more queries. However, observations showed that visually impaired searchers expressed their complete information needs in the form of long precise queries, and as a result, their queries were found to be more expressive. Sighted participants were found to place a strong emphasis on layout and aesthetics, while screen-reader users' impressions were thought to be largely dependent on content.

While prior work has offered an insight into the ways in which individuals who are blind explore and search for content, further work is needed to determine the ways in which credibility can be assessed. The study described in this paper, has aimed to identify the features on a web page used by individuals who are blind to assess credibility, and identify the browsing strategies/workarounds taken to make these assessments, compared with sighted peers. We have also aimed to examine the relationship between accessibility and credibility when exploring web content to make assessments. Insights from our study are thought to help to inform the design of highly credible interfaces for blind screen reader users.

3. OBSERVATIONAL STUDY

To investigate the ways in which credibility assessments are made, an observational study was undertaken. While the study primarily focused on the issues faced by individuals who are blind, sighted participants were also recruited in order to compare strategies/workarounds between groups.

3.1 Participants

Eleven legally-blind and eleven sighted volunteers were recruited for the study (aged 19 to 64: mean: 36). The snowball sampling technique was used, in order to identify blind participants with varying levels of experience using technology. All eleven legally-blind participants (B1-B11) had either limited or no residual vision, and relied upon screen readers to access content from the

Web (9 PC users favoring JAWS, 2 Mac users favoring VoiceOver). Six of the participants described themselves as congenitally blind, with the remaining five stating that they became blind in later life. Three had some level of light perception. The sighted group followed a similar age and sex distribution to the blind group.

Each of the participants described their level of web expertise on a scale of 1 to 5, where '1' related to being a 'novice user', while '5' related to being an 'advanced user'. Blind participants rated themselves as 4 on average, while sighted participants rated themselves as 4.55. Examples of tasks that advanced users mentioned they could perform independently included regular online shopping, purchasing tickets online, searching for information needed for purposes of college/work, and using social media.

All eleven blind participants expressed confidence in using their screen reader commands for browsing purposes (e.g., quick key navigation to more efficiently jump between different HTML elements on the page such as headings, links, and landmarks). They mentioned that the browsing process could be hindered if pages were designed without consideration for accessibility. For example, difficulties could be faced gaining an overview of content if pages were not designed with the appropriate HTML tags, or if graphics were not labeled.

3.2 Task design

Two tasks were designed which were presented to all participants:

- Task 1: Browsing 5 pairs of web pages examining a topic related to a specific query and then rating the credibility of these pages. *Example: Exploring two web pages which appear in the search results associated with the search term: "How to Reduce Personal Debt."*
- Task 2: Browsing 5 sets of search results present on a search engine results page (SERP), generated using a specific query, and then selecting the most credible result (i.e., which result would be most likely selected for purposes of exploration to satisfy the query). *Example: Assessing the search results found when searching for the term: "Renewable Energy."*

Stimuli for our study were selected from the dataset¹ generated by Schwarz and Morris [19]. The dataset contains 1,000 URLs and their corresponding cached web pages (covering five topics, with five queries per topic, and 40 search results per query), along with subjective topic expert credibility ratings for each URL. For the study described in this paper, the web pages selected covered five topics (celebrity news, environment, health, personal finance, and politics). Pages selected for both tasks varied in terms of credibility ratings assigned by topic experts in the study by Schwarz and Morris [19]. A further check was then performed by both authors/investigators from our paper (one using a screen reader) to independently rate the pages, and check for potential issues which may arise. For Task 2, as search result snippets were not present in the dataset used in Schwarz and Morris' study [19], these snippets were generated using a popular search engine. These were then ordered in a similar way to the original searches conducted in [19].

3.3 Running the study

The study was conducted both in-person and remotely. While all sighted participants were able to attend the testing venue, due to difficulties recruiting blind participants from the local area, ten

were asked to perform tasks remotely using video conferencing software.

Participants were provided with ten minutes of training, and then asked to perform both sets of tasks while thinking aloud. It was suggested that tasks should be completed as quickly as possible, without compromising quality. If the task could not be completed within a five minute period, participants were asked to move to the next task.

For Task 1, participants were asked to rate the credibility of each page examined and rate their confidence in assigning credibility ratings using Likert scales (1-5). Blind participants were also asked to rate the accessibility of each page. Participants were then asked to reflect upon their experience browsing each of these pages for Task 1, along with their experiences exploring the search results presented in Task 2. The post-task discussion helped us to clarify browsing strategies observed when performing the web-based tasks, as well as enabling participants to describe their reasoning behind their credibility assessment ratings.

Two investigators (one visually impaired, one sighted) were present for each session. Both took detailed notes. For purposes of analysis, both sets of notes were compared. Each session was audio recorded. The primary investigator (who identifies as visually impaired) listened to each recording, carefully examining the output from participants' screen readers when performing tasks. This step was taken to better understand the browsing strategies adopted, which might not have been explicitly verbalized during the sessions.

Table 1: Credibility ratings by user group, and the difference between these ratings with the corresponding values from the dataset used by Schwarz and Morris [19].

Web site	Credibility rating (mean)		Difference between ratings from our study compared with [19]	
	Sighted	Blind	Sighted	Blind
Michael Jackson-NYT	4.36	4.09	-0.64	-0.91
Michael Jackson- Flixster	3.09	2.55	0.09	-0.45
Organic-Heall	3.45	3.18	0.45	0.18
Organic-Grinning	2.73	4.00	-1.27	0.00
Autism-MedicineNet	4.27	4.27	-0.73	-0.73
Autism-AllExperts	2.73	2.70	-0.27	-0.30
Personal Debt-WiseGeek	2.27	3.09	0.27	1.09
Personal Debt-IdeaMarketters	2.36	2.89	0.36	0.89
Obama-CBS	4.36	4.18	-0.64	-0.82
Obama-FactCheck	4.09	3.55	-0.91	-1.45

4. RESULTS AND DISCUSSION

4.1 Quantitative findings

In order to analyze data from our study, the credibility ratings assigned to each web page selected for Task 1 were averaged for both blind and sighted groups. The difference between the average credibility rating for each page from our study and its corresponding rating from Schwarz and Morris' dataset [19], was then calculated. This step was taken to provide a point of

¹ <http://research.microsoft.com/credibility>

comparison (Table 1). For sighted participants, the greatest difference (-1.27) was identified for the Grinning page which related to organic eating (average rating of 2.73 vs. dataset rating of 4 – Figure 1). For blind participants, the greatest difference (-1.45) was identified for the FactCheck page (average rating of 3.55 vs. dataset rating of 5 – Figure 2). The features used to assess the credibility of these pages are described in 4.2.

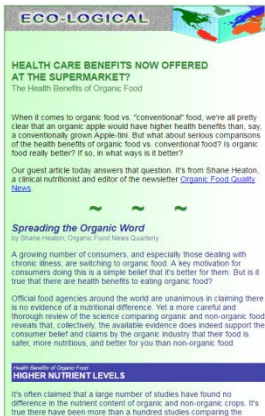


Figure 1: Grinning page

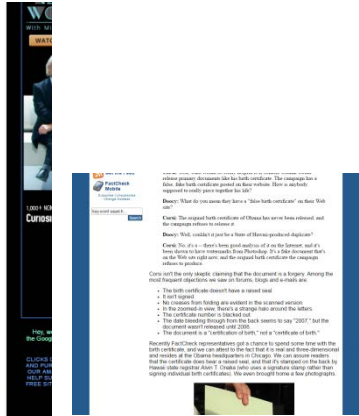


Figure 2: Fact Check page

The sighted group were found to assign ratings which were closer in value to the ratings from Schwarz and Morris’ dataset [19] (mean difference: 0.56/5), when compared with the blind group (mean difference: 0.68/5). Interestingly, the sighted group reported lower levels of confidence in assigning credibility ratings (4.28/5), compared with the blind group (4.36/5).

In order to examine the relationship between accessibility and credibility, a Spearman’s rank-order correlation was run on the ratings provided by blind participants (110 ratings from a total of 11 blind participants). Findings showed a positive correlation between both types of assessment, which was statistically significant ($r_s = .3352, n=110, p=.0003$). The correlation indicates a weak relationship between the accessibility of a web page and its perceived credibility.

4.2 Perceptions of credibility

Features influencing credibility varied between groups. Sighted participants identified using visual aesthetics and structural layout in order to assess credibility, similar to the findings of Fogg et al. [8], while blind participants focused more on the textual content. For example, one of the pages related to organic eating (Grinning) (Figure 1) was rated 2.73/5 by sighted participants. When questioned about their reasons for the rating, participants suggested that the page was not considered to be visually appealing, appearing somewhat outdated due to an “old-fashioned appearing color scheme”. The single column design was thought to be “fairly amateur-looking” in design. In contrast, the more “professional-looking” New York Times article (Michael Jackson-NYT) with a “clearer structure” and “more consistent formatting”, was rated higher (4.36/5). The New York Times was thought to be a brand that was recognizable and known for high standards in reporting, which contributed to the rating.

The FactCheck page (Figure 2) received a rating of 3.55/5 from blind participants. When questioned about their reasoning behind the rating, participants highlighted that the page lacked headings, and its main content appeared image-heavy. Although images were labeled, these were not as meaningful to the user, leading to difficulties gaining an overview of content. In contrast, sighted participants rated the page as more credible (4.09). They could

swiftly skim the textual content present on the page to gain an overview. The images were found to complement the text, aiding better comprehension of content.

4.3 Browsing Behavior

4.3.1 General browsing strategies

Browsing strategies were found to vary considerably between sighted and blind participants. Sighted participants were able to generally gain a near-instant overview of content and were able to traverse through longer pages with relative ease. Content could be skimmed fairly swiftly. The act of skimming enables users to glean more information about the ‘look and feel’ of a page (e.g., use of colors, typefaces and themes, visual structure of the page). Participants stated that features such as the presence and quality of images, and details relating to the author or reviewer of the content, would be helpful for making a more informed assessment of the page’s credibility.

In contrast, blind participants were observed following a more structured set of steps to gain an overview of content, with the goal of assessing credibility. Upon opening a new web page, participants were observed attempting to locate the main textual content present, to gauge the intended purpose of the page. To do this, a range of techniques were used, described in more detail in 4.3.2.

After locating and reading the main content, some of the blind participants were observed attempting to browse other sections of the page, in order to locate other features (e.g., menus or links to external resources) which would help lead to a more informed assessment of credibility. This attempt could have either a negative or positive impact on their perception of the page’s credibility, depending on how easily they could move around the page using their screen reader navigation shortcut keys (e.g., to navigate through headings, lists, etc.).

Blind participants stated that in order to browse pages containing considerable amounts of content (termed: ‘busy’ pages), reading speed and verbosity was adjusted to streamline the process of moving through information present. This would result in the participants expending considerable cognitive effort to listen attentively for certain terms or cues voiced by the screen reader, which would enable them to make a decision on how to proceed.

4.3.2 Non-visual browsing strategies

The following strategies were observed when blind participants were attempting to traverse content to make credibility assessments:

Avoiding extraneous content: Six blind participants were observed to use commands to skip over menu list items with the aim of moving straight to the main content present on the page. Participants who had accessed similarly structured pages in the past, described attempting to recall the steps previously taken to avoid extraneous information, including ads, list of links serving as menus, and social media links (e.g., Facebook, Twitter and Google Plus). On occasion, participants were noted to overshoot the main content. If this happened, participants would use commands such as the up arrow key to return back to the content that they had passed over.

Skipping graphics with inappropriate alternative descriptions: Images can be used for either decorative or informative purposes, when integrated with a web page. Six blind participants mentioned images without useful alternative descriptions would be bothersome and cognitively overloading when attempting to gain an overview of content, preventing them

from focusing on making an informed credibility assessment. B6 stated, “when I hear the term ‘graphic’, meaningless long numbers or things like ‘% &? 1287958445643 @’ outputted by my screen reader while browsing, I know that these images or frames are useless to spend time on, making me think that this content is not helping me in the least to determine the credibility of the page.” Participants were largely aware that sighted users could glance over these images easily, without wasting any time. However, they were willing to put in the time to traverse content sequentially, in order to reduce the risk of assessing the credibility poorly, particularly if information from the page would be used to support a decision.

Browsing sequentially not to miss information: Seven of our participants were observed browsing content line-by-line, even though it was more time-consuming than using other techniques. Since many of the pages presented were unfamiliar to them, participants mentioned that line-by-line browsing would help to reduce the likelihood of missing important pieces of information. This led to the participants feeling more confident that they had made fair assessments of a page’s credibility.

Using “HTML element” quick navigation: Participants were found to express greater levels of confidence when browsing familiar web pages, compared to ones that had not previously been accessed. Participants described developing mental structural representations of layout when exploring a page. While committing these representations to memory was found to be demanding, the technique could ‘pay-off’ if it helped to streamline the navigation process when visiting the site at a later time. Seven participants took advantage of their screen reader’s navigation features when exploring familiar pages. Actions performed included jumping between headings, buttons, or links present on pages. These could be used to gain an overview of content, on which they could form their decision about the credibility of the page. If satisfied, they would then switch back to sequential reading as their typical behavior to pay close attention to the content. B9 stated, “If I know the layout of the page, I mostly find the heading first to save time, then just press ‘say all’ command [‘Insert key plus Down arrow’] to read down.”

Reading text character-by-character: Four of the participants used this technique to read certain sections of a page, even though the process was time consuming in nature. This strategy was used to read URLs to determine credibility. Domain name extensions such as ‘.org’ were thought to include content that was more strongly vetted for presentation, compared with other extensions. Similar to Blythe et al. [2], participants described situations where they had selected URLs which were spelled or read by their screen readers in a similar way to commercial sites, but had inadvertently led them to unintended sites. Recovering from the error could be challenging, so it made sense to more cautious users to spend time going through the URL to avoid this from happening.

Searching for copyright and author information: In addition to checking URLs, four participants were observed checking for copyright symbols and digital certificates when exploring web pages. This technique has also been highlighted by Lazar et al. [14], as a method of assessing credibility. B1 was noted to spend time searching for the copyright symbol when exploring the Flixter page. After reading content present on the page, he was observed moving to the end of the page using the Ctrl and End keys. The up arrow key was then selected where he assumed using previous experience of accessing similar pages, that the copyright symbol would most likely be located. He stated “I know that

jumping up back to where I was before [in the text] can be a hassle.” He highlighted that it was worth making a trade-off, spending more time to locate features on a page which would inform a credibility assessment, rather than attempting to make a decision based on content from a page which may not be trustworthy.

Participants were found to also search for information about the page author(s). The presence of names along with credentials (e.g., MD, Ph.D.) would signify whether content had been written by someone qualified for the role. This would in turn impact the assessment of the page’s credibility.

4.4 Credibility assessment criteria

4.4.1 General characteristics of the page

Blind participants were observed focusing upon non-visual features present on web pages to support the assessment of credibility. These are described below.

Intention: Six blind participants attempted to identify the main message that each page was intending to convey (i.e., what was it selling, etc.) prior to providing a credibility assessment. They described searching for evidence in the textual content present on each page, to determine if the site appeared to be ‘professional-looking’ in nature, or whether it seemed more like a blog which may not have been vetted by a third party. The presence of ads, links to commercial products, and discussion related to the content, were criteria which used to determine the credibility of the page. As an example, while looking at the AllExperts web page which related to early signs of autism, B11 stated, “Oh, there are just questions and answers here. It seems that this page is nothing but just a blog post where advocates have posted responses.” As a result, the participant felt inclined to close the page and move to the next link.

‘Information Focus’ (term used by Fogg et al. [8]): Two participants (B5 and B6) closely examined whether content present on a page differed from the intended purpose of the site. For example, if participants detected links or discussions about topics other than the main topic associated with the page, it would raise doubts about the quality and validity of the content. For example, when browsing the Flixter page relating to Michael Jackson (MJ), B5 pointed out, “it seems that it is just a kind of media page with videos and images of many people. So it is not a specific page dedicated to MJ, making me doubtful about its credibility, compared to what I read about him on the New York Times page.” In contrast, sighted participants were not observed spending time attempting to differentiate between content on a site and its intended purpose.

4.4.2 Textual sensitivity

Textual content was found to be more heavily scrutinized by blind participants compared with the sighted group. The following highlights the textual characteristics which blind participants described they would use to help assess credibility.

Quality of the writing: Five participants stated that the quality of writing can provide a valuable insight into the accuracy of content. Depending on the context of the topic, this included how clearly the text was written, the sophistication of the writing (e.g., the use of technical terms), and how relevant the content was to the headings on the page. For example, looking at a blog-like web page with a casual style of writing, B11 pointed out, “The text does not appear to be expertly written when I skim it. It doesn’t pass my trust test.” Looking at the same page, B5 mentioned, “I cannot tolerate written text on a page that does not match its main purpose or topic.”

‘Tone of the writing’ (term described by Fogg et al. [8]): Three participants were found to skim the text using the ‘Say-all’ command, to establish whether the content was well constructed, and to identify the ‘tone’ of the writing. They pointed out if they felt the text was written in a casual style, especially in pages with a scientific topic, they would not rate the information to be as highly credible. Looking at a page which related to the early signs of autism, B11 mentioned, “It looks that there is just a kind of casual conversation going on here between someone who has asked a question and the responder. I will never take contents like this seriously. I prefer reading pages [relating to medical conditions] which are more professionally written, with scientific information present.”

Writing neutrality: Six blind participants stated when they skimmed through the main content of a web page, they would try to understand if the text was written as an opinion piece or a factual piece. If it was an opinion piece, participants described attempting to identify whether bias was present in the content. Although sighted participants mentioned this issue, they were found to mainly form their judgment based on whether or not the source appeared to be a neutral content provider (e.g., web sites with a reputation for presenting information with minimal bias), rather than spending time examining the content.

Writing mechanics: Grammatical errors and typos in the text negatively impacted the credibility ratings among four blind participants, while no sighted participants from our study explicitly referred to these. B5 pointed out, “...I can easily distinguish where there is a typo or grammatical error especially if I listen to the screen reader at a slower [reading] speed. You know it is quite clear through the change in tone of the spoken text. How can I trust the content and author when no-one has even proof read what they are presenting to me?”

4.4.3 Supporting Evidence as Criteria

The ability to locate supporting evidence in the main content of a page was found to impact credibility ratings among blind participants. Although sighted participants would take some of these factors into account, none of these were among the main criteria that they described as impacting their assessments.

Dates and statistics: When skimming text, three blind participants stated if they browsed a scientific or sensitive topic, they would pay close attention to statistical information present to support existing content, and any dates indicating when the page was developed, particularly when judging the credibility of time-sensitive topics (e.g., news articles, etc.). Participants were aware that older pages may contain information which may be out-of-date. Decisions based on out-of-date/inaccurate content were thought to lead to negative outcomes.

Presence of relevant links and citations: Seven blind participants stated that the presence of relevant links and citations would help to boost their perceptions of the page’s credibility. If they were not familiar with the topic associated with the page, links would be useful to identify more information about that topic. For example, when looking at a medical web site relating to autism (Medicine.net), B11 stated, “The page appears to be credible as it has links to provide an explanation of the terms that may be unfamiliar to its readers. This site seems to be more credible, as it supports the needs of users with limited knowledge of medical terms.” The page was noted not to contain links which were unrelated to the topic (autism).

Meaningfulness of the content and flow of the text: Four blind participants expected highly credible sites to be well structured,

with text that flows well. Detailed relevant content on the page also led to perceptions of higher credibility. B6 highlighted that the lack of information present on a page explaining issues related to autism (AllExperts site) made it difficult to determine whether much time had been spent developing the site. As a result, it was not thought to be as credible as the autism article on the Medicine.net site.

4.4.4 Familiarity with topic or site

Both blind and sighted participants used criteria such as personal expertise/prior knowledge about the topic, to support their judgments of credibility. For instance, B10, who defined himself as being knowledgeable about organic eating practices, stated while referring to the Heall page, “I am well familiar with this subject. When I read this page, it is obvious that there is some bias in the language.” This appeared to negatively impact his rating of credibility.

4.5 Impact of accessibility on perceptions of credibility

Blind participants emphatically stated that accessibility challenges when browsing web sites would not influence their perceptions of credibility. For example, B6 mentioned, “You know what? I know that people out there may not necessarily take accessibility into account when designing sites, but it does not mean what they present is not credible. I do my best to read the text and see what else I can find there to make an informed credibility assessment.” However, in practice, observations revealed that when accessibility barriers were encountered, ratings often appeared to be impacted. Instances are described in 4.5.1-4.5.3.

4.5.1 Page layout and inappropriate usage of visual UI elements

Inappropriate design was one challenge described by participants that could impact the process of skimming the main content. Three blind participants were observed to struggle traversing through the tables (used for layout purposes) embedded within the Grinning page relating to organic eating. These were used to visually organize objects and text on the page. Ironically, the layout of content was not appreciated by sighted participants (4.2). When asked to describe his experience accessing the page using a screen reader, B1 stated, “It is taking too much time to explore. I usually don’t care about pages that waste my time just to find content in them.” Although he mentioned that this influenced his credibility rating, he highlighted that there were positives which could help to offset the rating (e.g. the clarity of textual content present on the page).

When describing general experiences browsing sites, B1 stated that tags are often misused by interface designers, which can impact the accessibility of content. For example, the tag may be used by designers to embolden content when developing headings. However, heading tags (e.g., <h2>, <h3>), which make text appear visually similar to emboldened content, would be more useful, as screen readers are able to present lists of headings to the user to help gain an overview of content.

4.5.2 When content loaded slowly or was incomplete

The loading speed of the page content, though not directly an accessibility issue, was another factor that impacted credibility for four blind participants in our study. There were several instances when pages loaded more slowly than expected. Participants began to listen to the content present. However, only part of the information was present at this time as the page was still loading, which led to a negative perception of the site. Participants thought

that either the page was not finished by the author, or that part of the content was inaccessible. This was not observed as a problem for sighted participants because they were able to visually recognize whether or not the pages were loaded completely.

4.5.3 Inaccessible media objects

The presence of inaccessible media objects including Flash content, unlabeled buttons, or images without appropriate alternative descriptions negatively impacted the credibility rating for some pages. The impact was most obvious when the page was unfamiliar to the user. For example, when browsing a web page relating to a celebrity, which contained multiple images of the star, B8 stated, "I am not sure if these are his photos as there is no description that I can trust. Should I trust it when I cannot see it? Sighted users can see them, so it must be simpler for them to identify if the page is credible." In contrast, on pages where graphical objects which were informative in nature, and associated with alternative descriptions (e.g., on the page related to organic eating), participants reacted to the page content with a more positive view relating to its credibility.

4.6 Exploring search results

Blind participants in our study were generally confident with the process of conducting online searches, and were able to move quickly through the search results present in Task 2. Findings from Task 2 highlighted certain parallels between both groups of users when attempting to gain an overview of content. For example, sighted participants were able to skim through the search result titles, which were larger in size and blue in color, and therefore more noticeable on the SERP. Blind participants tended to use shortcut keys to move through the search results, skimming titles in a similar fashion to sighted users. As the search result titles generated by the Google site, are tagged with "Heading 3" tags, selecting either "H" or "3" keys would help a screen reader user to jump quickly between search result titles in sequence. After skimming through titles, both groups could focus on the longer snippets, to identify the relevance and associated credibility of the page.

Both blind and sighted participants described using similar subjective criteria to evaluate search results. For example, familiarity with the topic, source of the search result, and purpose of the search were taken into account when making credibility assessments. Both groups were also noted to place a certain amount of trust in the algorithms used by Google, Bing, and other popular search engines, which enabled them to feel more confident that results on the first page of the SERP might be more relevant, and likely credible. Neither group favored spending time moving to a second page of results, should they be unable to locate appropriate results from the first page. They favored reformulating the query instead.

5. INSIGHTS FROM THE STUDY

Researchers suggest that web users are becoming more skeptical of the information they find online and may be wary of web-based experiences [7]. The onus is placed upon interface designers to enhance credibility. While design guidance has often focused upon improving the visual appearance of a site, findings from our study reveal that to better meet the needs of individuals who are blind, other considerations need to be made.

- **Designers should be aware that perceptions of credibility differ when exploring content visually and non-visually. Attempts should be made to bridge this gap.** Sites which are more visually-appealing were found to be associated with greater levels of credibility among sighted participants (e.g.,

Flixtor page - Table 1). In contrast, sites with rich, well formulated textual content, and a strong structure, led to more favorable credibility ratings by blind participants, even though some of these pages appeared less visually attractive and therefore less credible to sighted participants (e.g., Grinning page). Interface designers should consider ways to ensure that all users, irrespective of ability, are able to obtain the cues they need to make an informed decision as to the credibility of content.

- **Designers should be aware of features which are important to assess credibility non-visually, and note that relative importance of features may vary between user groups.** Findings from our study highlighted that pages found to be highly credible by blind participants, stayed on topic (information focus), and contained limited extraneous information (e.g., adverts, links to unrelated topics). Credibility was negatively impacted by inappropriate tagging, and poor layout of content, both of which would make the process of navigating with a screen reader more challenging. In terms of textual content, the quality and 'tone' of the writing were found to influence ratings. While some of these features influenced ratings by sighted participants, visual aesthetics and structural layout were found to be the most important for purposes of making an assessment.
- **Sites should be designed taking into account the browsing strategies that are used by blind users.** In order to explore unfamiliar pages, techniques such as moving from heading-to-heading or link-to-link were often used to gain a quicker overview of page content, with a view to making credibility assessments. If a page is inappropriately coded, these strategies would be difficult to perform. Furthermore, information which may be helpful to sighted users may be extraneous to blind users (e.g., listening to the alternative text associated with non-informative images), and therefore may be skipped over to save time in this process. Pages which were more familiar to blind users (e.g., search result pages presented by a search engine which is frequently visited) could be explored with greater levels of confidence, removing an additional hurdle to making a credibility assessment.
- **Cautious users should be supported when making credibility assessments.** Some of the blind participants in our study favored reading content line-by-line or character-by-character, depending on whether important decisions rested upon whether the content presented was credible or not. Similar to [2], the screen reader could be used as a 'security device', enabling users to examine URLs which may be similar sounding to a commercial site's URL, but may lead to a different, or possibly malicious site. Blind participants felt that a greater time investment at this stage would lead to a more informed decision relating to credibility.
- **Design for accessibility, which in turn may influence credibility.** Findings from our study confirmed the presence of a positive correlation between accessibility and credibility. While blind participants in our study were hesitant to describe a link between the two factors, their browsing behavior indicated otherwise (see Section 4.5). Thorough accessibility testing is needed with individuals who are blind and visually impaired to reduce the risk of issues that may be faced when using a screen reader to browse a page. Additionally, identifying credibility and confidence in the

rating provided during the testing stage would also present considerable value.

6. CONCLUSION AND FUTURE WORK

The observational study described in this paper has examined the credibility assessment process undertaken by individuals who are blind and sighted. Two sets of web-based tasks were presented to participants. The features which are used to help inform credibility assessments have been identified. Furthermore, the relationship between accessibility and credibility has been described. Insights from the study can be used by interface designers when developing highly credible sites to cater to the needs of individuals who are blind and use screen readers to access the Web.

As the next logical step in the research, we aim to conduct a study where participants are presented with interfaces with varying levels of accessibility. Findings would offer a deeper insight into the ways in which inappropriate design can impact both accessibility and credibility assessments of a site. Further work may also be conducted to examine the ways in which assessments of credibility may vary when the situation, context, or environment differs (e.g., investigating search results while on-the-go using mobile devices). Findings would help to develop and strengthen guidance for interface designers aiming to support levels of credibility among users with diverse needs and abilities.

7. ACKNOWLEDGMENTS

We would like to thank Leila Heroabadi, Dhuel Fisher and William Easley, who helped prepare study materials and run study sessions; Josanne Revoir, Stacy Branham and Flynn Wolf, who helped proof-read the document; and Erick Ronquillo and Nasrin Attaran, who helped the visually impaired author format the manuscript. We also thank our participants for their valuable feedback.

8. REFERENCES

- [1] Andronico, P., Buzzi, M., Castillo, C. and Leporini, B. 2006. Improving search engine interfaces for blind users: a case study. *Universal Access in the Information Society*. 5, 1 (2006), 23–40. DOI= <http://dx.doi.org/10.1007/s10209-006-0022-3>
- [2] Blythe, M., Petrie, H. and Clark, J.A. 2011. F for fake: four studies on how we fall for phish. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI'11. ACM, New York, NY, 3469–3478. DOI= <http://doi.acm.org/10.1145/1978942.1979459>
- [3] Borodin, Y., Bigham, J.P., Dausch, G. and Ramakrishnan, I.V. 2010. More than meets the eye: a survey of screen-reader browsing strategies. In *Proceedings of the International Cross Disciplinary Conference on Web Accessibility*. W4A'10. ACM, New York, NY, Article No. 13. DOI= <http://doi.acm.org/10.1145/1805986.1806005>
- [4] Craven, J. and Brophy, P. 2003. *Non-Visual Access to the Digital Library (NoVA): the Use of the Digital Library Interfaces by Blind and Visually Impaired People*. Technical Report. Centre for Research in Library & Information Management, The Manchester Metropolitan University.
- [5] Fogg, B., Marshall, J., Kameda, T., Solomon, J., Rangnekar, A., Boyd, J. and Brown, B. 2001. Web credibility research: a method for online experiments and early study results. *Extended Abstracts on Human Factors in Computing Systems*. CHI'01. ACM, New York, NY, 295–296. DOI= <http://doi.acm.org/10.1145/634067.634242>
- [6] Fogg, B., Marshall, J., Osipovich, A., Varma, C., Laraki, O., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P. and Treinen, M. 2000. Elements that affect web credibility: early results from a self-report study. *Extended Abstracts on Human Factors in Computing Systems*. CHI'00. ACM, New York, NY, 287–288. DOI= <http://doi.acm.org/10.1145/633292.633460>
- [7] Fogg, B.J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P. and Treinen, M. 2001. What makes web sites credible?: a report on a large quantitative study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI'01. ACM, New York, NY, 61–68. DOI= <http://doi.acm.org/10.1145/365024.365037>
- [8] Fogg, B.J., Soohoo, C., Danielson, D.R., Marable, L., Stanford, J. and Tauber, E.R. 2003. How do users evaluate the credibility of web sites?: a study with over 2,500 participants. In *Proceedings of the Conference on Designing for User Experiences*. DUX'03. ACM, New York, NY, 1–15. DOI= <http://doi.acm.org/10.1145/997078.997097>
- [9] Freeman, K.S. and Spyridakis, J.H. 2004. An examination of factors that affect the credibility of online health information. *Technical Communication*. 51, 2 (2004), 239–263.
- [10] Hilligoss, B. and Rieh, S.Y. 2008. Developing a unifying framework of credibility assessment: construct, heuristics, and interaction in context. *Information Processing & Management*. 44, 4 (Jul. 2008), 1467–1484. DOI= <http://dx.doi.org/10.1016/j.ipm.2007.10.001>
- [11] Hovland, C.I., Janis, I.L. and Kelley, H.H. 1953. *Communication and persuasion; psychological studies of opinion change*. Yale University Press, New Haven, CT.
- [12] Ivory, M.Y., Yu, S. and Gronemyer, K. 2004. Search result exploration: a preliminary study of blind and sighted users' decision making and performance. *Extended Abstracts on Human Factors in Computing Systems*. CHI'04, ACM, New York, NY, 1453–1456. DOI= <http://doi.acm.org/10.1145/985921.986088>
- [13] Kim, J. and Moon, J.Y. 1998. Designing towards emotional usability in customer interfaces—trustworthiness of cyber-banking system interfaces. *Interacting with Computers*. 10, 1 (Mar. 1998), 1–29. DOI= [http://dx.doi.org/10.1016/S0953-5438\(97\)00037-4](http://dx.doi.org/10.1016/S0953-5438(97)00037-4)
- [14] Lazar, J., Meiselwitz, G. and Feng, J. 2007. Understanding web credibility: a synthesis of the research literature. *Foundations and Trends® in Human-Computer Interaction*. 1, 2 (2007), 139–202. DOI= <http://dx.doi.org/10.1561/1100000007>
- [15] Metzger, M.J. 2007. Making sense of credibility on the Web: models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*. 58, 13 (Nov. 2007), 2078–2091. DOI= <http://dx.doi.org/10.1002/asi.20672>
- [16] Murphy, E., Kuber, R., McAllister, G., Strain, P. and Yu, W. 2008. An empirical investigation into the difficulties experienced by visually impaired Internet users. *Universal Access in the Information Society*. 7, 1–2 (Apr. 2008), 79–91. DOI= <http://dx.doi.org/10.1007/s10209-007-0098-4>
- [17] Robins, D. and Holmes, J. 2008. Aesthetics and credibility in web site design. *Information Processing & Management*. 44,

1 (Jan. 2008), 386–399. DOI=
<http://dx.doi.org/10.1016/j.ipm.2007.02.003>

[18] Sahib, N.G., Tombros, A. and Stockman, T. 2012. A comparative analysis of the information-seeking behavior of visually impaired and sighted searchers. *Journal of the American Society for Information Science and Technology*. 63, 2 (Feb. 2012), 377–391. DOI=
<http://dx.doi.org/10.1002/asi.21696>

[19] Schwarz, J. and Morris, M. 2011. Augmenting web pages and search results to support credibility assessment. In

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI'11. ACM, New York, NY, 1245–1254.

DOI=<http://doi.acm.org/10.1145/1978942.1979127>

[20] Wathen, C.N. and Burkell, J. 2002. Believe it or not: factors influencing credibility on the web. *Journal of the American Society for Information Science and Technology*. 53, 2 (2002), 134–144. DOI= <http://dx.doi.org/10.1002/asi.10016>