



Assessing Google Search's New Features in Supporting Credibility judgments of Unknown Websites

Ace Wang

aw117@wellesley.edu
Wellesley College
Wellesley, Massachusetts, USA

Liz Maylin De Jesus Sanchez

ld102@wellesley.edu
Wellesley College
Wellesley, Massachusetts, USA

Anya Wintner

aw105@wellesley.edu
Wellesley College
Wellesley, Massachusetts, USA

Yuanxin Zhuo

yz101@wellesley.edu
Wellesley College
Wellesley, Massachusetts, USA

Eni Mustafaraj

emustafa@wellesley.edu
Wellesley College
Wellesley, Massachusetts, USA

ABSTRACT

This study assesses the awareness and perceived utility of two features Google Search introduced in February 2021: "About this result" and "More about this page". Google stated that the goal of these features is to help users vet unfamiliar web domains (or sources). We investigated whether the features were sufficiently prominent to be detected by frequent users of Google Search, and their perceived utility for making credibility judgments of sources, in one-on-one user studies with 25 undergraduate college students, who identify as frequent users of Google Search. Our results indicate a lack of adoption or awareness of these features by our participants and neutral-positive perceptions of their utility in evaluating web sources. We also examined the perceived usefulness of nine other domain credibility signals collected from the W3C.

KEYWORDS

user study, web source evaluation, evaluation of search results, credibility signals.

ACM Reference Format:

Ace Wang, Liz Maylin De Jesus Sanchez, Anya Wintner, Yuanxin Zhuo, and Eni Mustafaraj. 2023. Assessing Google Search's New Features in Supporting Credibility judgments of Unknown Websites. In *ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '23)*, March 19–23, 2023, Austin, TX, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3576840.3578277>

1 INTRODUCTION

Google introduced the "About this result" and "More about this page" features in February 2021 at a time when misinformation, particularly related to COVID-19 and vaccines, was spreading rapidly online. The "About this result" panel is accessed by clicking the vertical ellipsis next to each result's URL. Within this panel is a link

to "More about this page" (see Figure 1 and footnote¹). Google's goal with these features is to make it easier for people to vet unfamiliar sources and to provide additional assurance and peace of mind when searching. The example provided in their press release demonstrated how the features can be used to verify the credibility of information from a medical journal, The Lancet, which is unfamiliar to many outside the medical research community [11]. These features let individuals quickly access what other websites have to say about a source or topic, a strategy known as "lateral reading", recommended by information literacy researchers [2, 3, 20]. While we believe Google's new features have the potential to be a helpful tool for people, we wondered whether Google's vertical ellipsis icon that brings up the features would be sufficiently prominent to be engaged with. Prominence-Interpretation Theory posits that successful interventions must be both prominent enough to draw attention and also be interpreted correctly by the user [4]. Would Google's new features and the vertical ellipsis to access these features be prominent enough to draw attention to regular Google Search users or go unnoticed? Furthermore, would the information provided in these features be interpretable and useful for search users? To the researchers' best knowledge, there have been no publicly available user studies published to date on Google's "About this result" and "More about this page" features.

We conducted 25 one-on-one interviews wherein we executed a think-aloud procedure, collected survey responses, and observed search behaviors. We found zero prior awareness or use of the two features when participants were given their initial task of finding credible sources via search to help a friend in a toxic relationship. When shown these features after their search task, participants found the features no more than "somewhat helpful". Some participants also found the "More about this page" feature confusing and were unclear of its purpose in supporting lateral reading.

2 RELATED WORK

2.1 Critical Thinking on Search

In web search, the heuristic of "high ranking equals high-quality content" has been found again and again to have the biggest impact on click-through behavior and credibility judgments [8, 14]. This rank order bias persists even when the top result is unrelated to

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 the author(s) 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.
CHIIR '23, March 19–23, 2023, Austin, TX, USA

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0035-4/23/03...\$15.00
<https://doi.org/10.1145/3576840.3578277>

¹Due to space limitations, high-resolution screenshots are available online at this page: <https://cs.wellesley.edu/~credlab/chiir23/>.

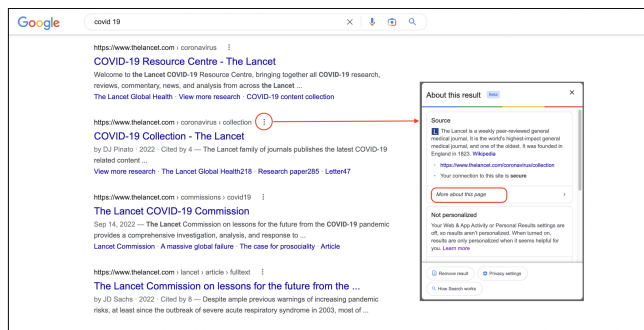


Figure 1: Google’s “About this result” and “More about this page”.

the search task [6, 10, 16, 19]. This suggests that users trust Google to present the most relevant and high-quality results at the top of the search engine results page (SERP), and may not engage active critical thinking when searching. Organic search results typically include only a title, short snippet, and URL. This may not provide enough information for users to accurately assess the quality or credibility of the content. Information Foraging Theory [15] and researchers such as Schwarz and Yamamoto [17, 23] have proposed that this strong rank order bias may be due to the limited information available on the page to guide users’ decisions. Smith and Rieh have suggested designing search systems that provide more knowledge-context and meta-information to search users. They propose a design vision for search systems that not only optimize for accuracy and efficiency in answer retrieval and task completion but also prioritize long-term learning and promote information-literate actions. They suggest that providing knowledge-context and meta-information about sources on SERPs can facilitate web literacy skills such as comparing, evaluating, and differentiating between information sources without hindering efficient task completion [18]. HCI researchers are increasingly studying the effectiveness of nudges and signals on SERPs. However, many disclosures and signals go unnoticed or ignored [14, 22]. According to Fogg’s Prominence-Interpretation Theory, interventions like signals and nudges can only be successful if they meet the conditions of being both prominent enough to draw attention and interpreted by the user as important and salient [4]. We wanted to understand if Google’s features would be prominent and interpretable enough to be effective in carrying out the Google’s stated mission of helping users vet unfamiliar sources with these features.

2.2 Augmenting SERPs with Credibility Signals

There is active research in the area of augmenting SERPs with credibility signals to improve users’ assessments of content and sources. For example, Schwaz and Morris tested the efficacy of displaying domain popularity among experts in the field and the general public [17]. Winter et al. studied the impacts of displaying the author’s name and profession and a community rating of the author [21]. One major challenge of implementing successful content signals is accounting for variations in cognition, expertise and motivation, as was demonstrated by Kattenbeck. He found that different people interpreted the same content cues differently [9]. Google is among the many content

platforms utilizing content labeling as a strategy for misinformation intervention. While the popularity of content labeling may be due to its perception as being less-divisive and political than fact-checks, researchers studying the efficacy of content labeling have also found that it can be more effective at shifting news consumption behaviors and attitudes towards sources than fact-checks alone [1]. Another major challenge for the approach of content labeling is identify signals that are both human-understandable and machine-operable. Many of the content labels in research studies are manually collected by researchers and may be challenging to implement at a wide-scale. The World Wide Web Consortium’s (W3C) *credible web community group*, which aims to shift the web towards more trustworthy content, drafted a report detailing characteristics of “promising” credibility signals. These characteristics include being difficult to manipulate, widely available, scalable, and verifiable. Of the hundreds of crowd-sourced credibility signals suggested, the W3C group identified Date Website First Archived, Corrections Policy, and Awards (e.g. Pulitzer Prize, RNG awards) as “promising” for research and development due to their characteristics [5]. In a previous study, Lurie and Mustafaraj examined Google’s Knowledge Panel and “Awards” tab in assessing the credibility of news sources [12]. They found that the “Awards” tab, which displayed wins and nominations for Pulitzer Prizes, was the most valuable component of Google’s Knowledge Panel for participants in evaluating sources. However, the news-source specific knowledge panel no longer exists, and it has been speculated that Google removed it due to political pressure [13]. Google’s new “About this result” feature contains information about the first archival date of the website but does not present any awards information. In the latter part of our study, we asked participants to evaluate the perceived utility of including awards and some of the other credibility signals that the W3C credible web community group crowd-sourced.

3 DATA AND METHODS

3.1 Participants

We conducted 25 interview based user sessions with students at a liberal arts college who were given a \$15 Amazon gift card for their participation. All participants identified as frequent users of Google Search and felt confident in their search abilities. We identified college students as a user group that could benefit from these new Google features and chose them as our study subjects due to the frequency of their interactions with search engines and the frequency of which they utilize Google Search to identify credible information for school work. Prior studies have found that college students often have difficulty identifying reliable information online and evaluating the credibility of sources, often relying on superficial criteria such as the website’s design or the presence of images to evaluate the credibility of sources [7].

3.2 Stimuli

Prior to our user study, 130 user-generated search queries related to toxic relationships were collected through interviews with sixteen undergraduate students (n=16). For each of these 130 queries, the first page of these search engine results pages (SERPs) was captured

and analyzed with Tranco², a research tool that analyzes site rankings, to identify a SERP with a mix of websites of varying credibility levels. The SERP for the query “How to learn to receive healthy love” was identified as having a mix of high-popularity websites and less-known websites of varying credibility levels. Two of the authors manually analyzed the domains on this SERP using credibility criteria from Google’s Search Rater Guidelines, W3C signals, and lateral reading to validate this mix of high and low credibility domains on this SERP. As Google search results and their rankings can regularly change, we created a simulated clickable SERP for the chosen query. This allowed the researchers to ensure that each participant was exposed to the same set of search results and that researchers could easily track the elements that participants clicked on. We used the interface design tool Figma³ to create a clickable mock-up and imported it into Useberry⁴ to collect SERP interaction data such as click-maps. Our simulated environment had the same clickable elements as the original SERP on Google. Participants’ “think-aloud” and ratings were collected through forms filled out by the participants and the interviewer.

3.3 Procedure

The study sessions took place from 07/15-07/20/2022 and lasted 20-30 minutes. The study was approved by our institution’s IRB and all participants gave verbal consent.

In one-on-one facilitated think-aloud sessions, participants interacted with the simulated SERP to complete the task of finding information to help a friend in a toxic relationship. Their search behaviors, including clicks and scrolling, were recorded by Useberry and they were asked to think-aloud as they worked through the task. After 3 minutes, the researchers asked participants about their familiarity with the sources on this SERP and to explain how they assess the credibility of unfamiliar sources. Subsequently, researchers asked participants about their familiarity with the “About this result” panel and what they expected from this feature. We recorded no interaction with Google’s new features by study participants as they completed the search task. After the initial task and set of interview questions, participants were given a chance to interact with the “About this result” feature and were asked to rate it on a scale of 1-5 in usefulness in vetting unfamiliar sources. This procedure was repeated with the “More about this page” feature. Participants were then given a description of nine other credibility signals and asked to rate them on a scale of 1-5 in terms of usefulness in vetting sources. These signals - Transparency, Content-Type, Publication Type, Web Accessibility, Authorship, Advertising, Awards, Originality, and References - were selected from a larger list of domain credibility signals from the W3C web credibility group [5].

4 FINDINGS

4.1 Awareness and Perceived Utility of Google’s new Features

The “About this result” and “More about this page” features had been available on Google’s search results pages for over a year and a half, but none of the participants (n=0) were aware of them despite

being regular Google users. Once familiarized with the “About this result” feature, our participants had a neutral-positive reaction (mean 3.52, standard deviation 1.046). When asked to explain their ratings, several participants commented that while the feature could help them assess sources, it is not sufficiently helpful to use in daily life and they can easily conduct another google search to gather information about the credibility of a site. Figure 2 shows a summary of the elements that participants found most useful in the “About this result feature”, though none of the individual elements received a majority of participants’ votes for usefulness.

Useful elements in “About this result” panel	Description	Votes (out of 25)
Background information	Description excerpted from Wikipedia (only for well-established sources).	10
Relevance	Relevant keywords from search query and why this source was presented as a result.	10
Security	Whether the domain provides a secure internet connection	6
Index information	Date that the domain was first indexed by Google.	2

Figure 2: Most useful elements of Google’s “More about this page” feature according to study participants.

Useful elements in “More about this page” feature	Description	Votes (out of 25)
Self description	A description quoting from the domain’s “About Us” page.	11
Google description	Description of the domain from wikipedia or other reference site when applicable.	7
More content from the domain	Shows users web pages from the same domain.	6
Related article	Shows users whether their connection to the site is secure or not.	4
Web results about the source	Shows users web results about the source, usually from third-party sites.	2

Figure 3: Most useful elements of Google’s “More about this page” feature according to study participants.

For the “More about this page” feature, we also found neutral-positive ratings with 70% of participants rating it neutral to somewhat useful (3-5 on a 5-point scale), with a mean of 3.28 and a standard deviation of 1.021. We show the most voted elements in Figure 3. Again, none of the elements were chosen by most participants, who in general were more critical in the open-ended comments during the think-aloud. Participants said that the feature felt very similar to the “About this result”, as they had expected to see more information such as date published, authorship, ratings, or other indications of credibility. They also noted that certain sites only included the web owner’s self review rather than reviews from others. This primarily occurred for less established domains, which was frustrating, as less recognizable domains are where participants sought the most help from Google’s features. Some participants voiced that it was easier to perform a second search rather than use the “More about this page” feature, undermining its purpose in assisting with lateral reading.

²<https://tranco-list.eu/>

³<https://www.figma.com/>

⁴<https://www.useberry.com>

4.2 Suggestions of Additional Credibility Signals

Before introducing W3C's nine credibility signals to our participants, they were asked for suggestions on how to make the two existing Google features more helpful. Many suggested third-party assessments like citations, reviews, site traffic ($n=8$), information on the author ($n=7$), publication funding sources ($n=5$), and various other disclaimers such as awards, ads, and plagiarism. This was satisfying to see, since many of these signals matched the researchers' list of proposed W3C credibility signals. We have summarized participant reactions to each of the proposed signals in Figure 4, where a score of 4 or 5 on a 5-point scale indicates that the signal would be helpful.

Signal	Description	Votes for helpful (Out of 25)
Page-Level Credibility Signals		
References	Citations that support domain claims and demonstrate prior research.	23
Authorship	The author name listed for the content, or another description like anonymous.	22
Content Type	Category of the published content.	18
Originality	Unique content versus modified or re-used content.	14
Domain-Level Credibility Signals		
Publication Type/Owner	This signal aims to show who is responsible for the site, content, or article.	23
Transparency	Presence of an About and/or Contact page.	18
Web Accessibility	Features that prevent people with disabilities or low internet bandwidth from accessing content and tools.	16
Advertising	Number of advertisements on page.	15
Awards	Indicates the awards a domain has received.	13

Figure 4: Descriptions of 9 W3C credibility signals and participant rankings as either “somewhat useful” or “completely useful”.

4.3 Study Limitations

Our study was limited in the number and diversity of participants, 25 undergraduates at one liberal arts college. Extensions of this study could survey larger groups and a more diverse population of Google search users. We hope to build upon this current study with a larger-scale study with a greater diversity of Google Search users to understand if our results can be generalized to non-college Google users.

5 CONCLUSION

In our small scale user study, we found that all participants lacked familiarity with the “About this result” and “More about this page” features of Google Search. Secondly, when introduced to this feature, most participants felt neutral toward the information that Google provided. A few participants felt confused about why Google would provide lateral reading sources on the “More about this site” page. Many participants felt that the features did not contain sufficient information to be useful and adopted in their regular everyday search habits.

Following Fogg's P-I Theory, we believe there are opportunities to improve both the prominence of the “About this result” and “More about this page” features and to choose an icon that is more easily interpreted by users as containing source information. Not only is the current vertical ellipsis easy to miss when conducting a search, it is also unclear from the icon what information will appear upon click.

Secondly, we believe there are opportunities to improve the signals contained within the features themselves. While content labels can help users by providing additional knowledge-context and foster critical thinking on the validity of information encountered on the Web, labels can also contribute to information overload, become white noise and simply ignored. The knowledge-context provided is currently limited to connection security, archival date, and wikipedia description in the “About this result” feature. When we asked participants to evaluate nine other credibility signals, we found that site ownership, authorship, and references by other sites were deemed by participants to be high utility in vetting unknown sources.

With web users relying heavily on Google Search to make life decisions big and small, from what to purchase to who to elect, the issue of being able to trust what we read online becomes ever more critical. We believe that useful credibility signals are important to develop, as misinformation spread faster than platforms' and fact-checkers' ability to counter it. Additional research can be conducted to assess various augmentations of signals and labels on SERPs as well as the efficacy of current platform interventions on helping users evaluate content and sources.

ACKNOWLEDGMENTS

We are grateful to the members of the Wellesley Cred Lab for supporting this research in various ways, and acknowledge funding from the National Science Foundation, under grant IIS 1751087.

REFERENCES

- [1] Kevin Aslett, Andrew M. Guess, Richard Bonneau, Jonathan Nagler, and Joshua A. Tucker. 2022. News credibility labels have limited average effects on news diet quality and fail to reduce misperceptions. *Science Advances* 8, 18 (May 2022). <https://doi.org/10.1126/sciadv.abl3844>
- [2] Joel Breakstone, Mark Smith, Priscilla Connors, Teresa Ortega, Darby Kerr, and Sam Wineburg. 2021. Lateral Reading: College Students Learn to Critically Evaluate Internet Sources in an Online Course. *Harvard Kennedy School Misinformation Review* 2, 1 (Feb 2021). <https://doi.org/10.37016/mr-2020-56>
- [3] Mike Caulfield. 2017. *Web Literacy for Student Fact-Checkers*. Washington State University Vancouver, New York, NY. <https://digitalcommons.liberty.edu/textbooks/5>
- [4] B. J. Fogg. 2003. Prominence-Interpretation Theory: Explaining How People Assess Credibility Online. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) (CHI EA '03). Association for Computing Machinery, 722–723. <https://doi.org/10.1145/765891.765951>
- [5] W3C Community Group. 2020. Reviewed Credibility Signals W3C Community Group Draft Report. <https://credweb.org/reviewed-signals-20200224>. Accessed: 2022-10-26.
- [6] Alexander Haas and Julian Unkel. 2017. Ranking versus reputation: perception and effects of search result credibility. *Behavior and Information Technology* 36, 12 (Dec 2017). <https://doi.org/10.1080/0144929X.2017.1381166>
- [7] Alison Head and Michael Eisenberg. 2010. Truth be told: How college students evaluate and use information in the digital age. *SSRN* (Oct 2010). <https://doi.org/10.2139/ssrn.2281485>
- [8] Yvonne Kammerer and Peter Gerjets. 2014. The role of search result position and source trustworthiness in the selection of web search results when using a list or grid interface. *International Journal of Human-Computer Interaction* 30, 3 (2014), 177–191 pages. <https://doi.org/10.1080/10447318.2013.846790>
- [9] Markus Kattenbeck and David Elweiler. 2019. Understanding credibility judgments for web search snippets. *Aslib Journal of Information Management* 71, 3 (March 2019), 368–391 pages. <https://doi.org/10.1108/AJIM-07-2018-0181>

- [10] Mark Keane, Maeve O'Brien, and Barry Smyth. 2008. Are people biased in their use of search engines? *Commun. ACM* 52, 2 (Feb. 2008), 49–52 pages. <https://doi.org/10.1145/1314215.1314224>
- [11] JK Kearns. 2021. A quick way to learn more about your search results. <https://blog.google/products/search/about-search-results/>. Accessed: 2022-10-26.
- [12] Emma Lurie and Eni Mustafaraj. 2018. Investigating the Effects of Google's Search Engine Result Page in Evaluating the Credibility of Online News Sources. In *Proceedings of the 10th ACM Conference on Web Science* (Amsterdam, Netherlands) (*WebSci '18*). Association for Computing Machinery, New York, NY, USA, 107–116. <https://doi.org/10.1145/3201064.3201095>
- [13] Emma Lurie and Eni Mustafaraj. 2020. Highly Partisan and Blatantly Wrong: Analyzing News Publishers' Critiques of Google's Reviewed Claims. In *Proceedings of the 2020 Truth and Trust Online Conference* (Virtual) (*TTO '20*). 64–72. <https://truthandtrustonline.com/wp-content/uploads/2020/10/TTO07.pdf>
- [14] Bing Pan, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gay, and Laura Granka. 2007. In Google We Trust: Users' Decisions on Rank, Position and Relevance. *Journal of Computer-Mediated Communications* 12, 3 (Apr 2007), 801–823 pages. <https://doi.org/10.1111/j.1083-6101.2007.00351.x>
- [15] Peter Pirolli and Stuart Card. 1999. Information Foraging. *Psychological Review* 106, 4 (Jan. 1999), 643–675 pages. <https://doi.org/10.1037/0033-295X.106.4.643>
- [16] Ladislao Salmeron, Yvonne Jammerer, and Pilar Garcia-Carrion. 2013. Are people biased in their use of search engines? *Computers in Human Behavior* 29, 6 (Nov. 2013), 2161–2171 pages. <https://doi.org/10.1016/j.chb.2013.04.034>
- [17] Julia Schwarz and Meredith Ringle-Morris. 2011. Augmenting Web Pages and Search Results to Support Credibility Assessment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) (*CHI '11*). Association for Computing Machinery, New York, NY, USA, 1245–1254. <https://doi.org/10.1145/1978942.1979127>
- [18] Catherine L. Smith and Soo Young Rieh. 2019. Knowledge-Context in Search Systems: Toward Information-Literate Actions. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval* (Glasgow, Scotland UK) (*CHIIR '19*). Association for Computing Machinery, New York, NY, USA, 55–62. <https://doi.org/10.1145/3295750.3298940>
- [19] Julian Unkel and Alexander Haas. 2017. The effects of credibility cues on the selection of search engine results. *Journal of the association for information science and technology* 68, 8 (June 2017), 1850–1862 pages. <https://doi.org/10.1002/asi.23820>
- [20] Sam Wineburg and Sarah McGrew. 2017. Lateral Reading: Reading Less and Learning More then Evaluating Digital Information. *SSRN* 121, 11 (Oct 2017), 56 pages. <https://doi.org/10.1177/016146811912101102>
- [21] Stephan Winter and Nicole C. Kramer. 2014. A question of credibility - effects of source cues and recommendations on information selection on news sites and blogs. *Communications* (Nov. 2014). <https://doi.org/10.1515/commun-2014-0020>
- [22] Min Wu, Robert C. Miller, and Simson L. Garfinkel. 2006. Do Security Toolbars Actually Prevent Phishing Attacks?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montreal, Quebec, Canada) (*CHI '06*). Association for Computing Machinery, 601–610. <https://doi.org/10.1145/1124772.1124863>
- [23] Yusuke Yamamoto and Katsumi Tanaka. 2011. Enhancing Credibility Judgment of Web Search Results. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) (*CHI '11*). Association for Computing Machinery, New York, NY, USA, 1235–1244. <https://doi.org/10.1145/1978942.1979126>