Visual Cues and Trust in AI-Generated Search Results: An Eye-Tracking Study on Google AI Overviews

Brüning Genannt Wolter Marius[a], Haase Lisa Eleonora[a], Lewandowski Dirk[a], Michaelsen Benno[a] and Schott Fabian[a]

[a] *Hochschule für Angewandte Wissenschaften Hamburg (HAW Hamburg), Berliner Tor 5, Hamburg, 20099, Germany*

1. Abstract

The integration of generative artificial intelligence (AI) into search engines is transforming how users access and evaluate online information. Search engine results pages (SERPs) have evolved from ranked lists of hyperlinks into dynamic, query-responsive interfaces that blend retrieved content with model-generated text (Amer & Elboghdadly, 2024; Trippas et al., 2025). Despite high user trust in search engines as information sources (Ries et al., 2025), AI-generated answers remain susceptible to factual errors and hallucinations, typically delivered with high confidence and accompanied by source attributions that frequently fail to substantiate the claims they accompany (Liu et al., 2023; Memon & West, 2024). Research consistently shows that users rely on heuristic cues, such as layout, design, branding, and citation presence, rather than verifiable accuracy when judging credibility (Fogg et al., 2003; Metzger & Flanagin, 2013; Ding et al., 2025). Visual trust signals such as logos or inline citations can increase perceived credibility independently of factual quality (Ding et al., 2025; Kang et al., 2015) yet may simultaneously obscure inaccuracies. Eye-tracking studies further demonstrate that even minor changes in SERP design significantly influence users' visual attention and behaviour (Lewandowski & Kammerer, 2021; Strzelecki, 2020).

In Germany, trust in journalistic brands varies considerably: public service broadcasters and established regional newspapers rank among the most trusted news sources, whereas tabloid and online-native outlets receive markedly lower trust ratings (Newman et al., 2025). This pre-existing differentiation in brand trust provides an ecologically valid basis for manipulating source credibility as an experimental factor, allowing the present study to examine how citation presentation, source credibility, and answer accuracy jointly shape user trust and content adoption on AI-generated SERPs.

To investigate these questions, a 2×2×3 mixed-factorial eye-tracking experiment was conducted with 43 participants in a usability laboratory in Hamburg, Germany. This design involved three manipulated variables (two of two levels, one of three levels), combined between- and within-participants comparisons, and included continuous eye-tracking throughout. Participants completed six health-related search tasks on simulated Google SERPs containing manipulated AI-generated answers. Independent variables were source presentation (logo versus footnote), source credibility (trusted versus less trusted), and answer accuracy (accurate, vague, false). The dependent variable was behavioural adoption of AI-generated content, defined as participants' subsequent reproduction of AI-provided answers on a follow-up questionnaire and multiple gaze-based measures, including the number, duration, and sequence of fixations on the manipulated SERPs across predefined areas of interest (AOIs).

Answer accuracy was the strongest predictor of adoption: participants were significantly more likely to adopt accurate answers than vague or false ones ($p < .001$). Neither source credibility ($p = .688$) nor

SEASON 2026, September 15–17, 2026, Hamburg, Germany
EMAIL: Marius.BrueningGenanntWolter@haw-hamburg.de (A. 1); Lisa.Haase@haw-hamburg.de (A. 2); dirk.lewandowski@haw-hamburg.de (A. 3); Benno.Michaelsen@haw-hamburg.de (A. 4); Fabian.Schott@haw-hamburg.de (A. 5);
ORCID: 0009-0009-2353-3030 (A. 1); 0009-0008-2852-7054 (A. 2); 0000-0002-2674-9509 (A. 3); 0009-0000-8004-3526 (A. 4); 0009-0007-9812-734X (A. 5);
DOI: https://doi.org/10.48441/4427.3720

visual presentation (p = .404) significantly influenced adoption, and no interaction effects were observed. Eye-tracking revealed that logos attracted significantly more fixations to the citation area than footnotes (p = .023), but this heightened attention did not translate into higher adoption or self-reported trust. Descriptively, false answers were adopted more often than vague ones, suggesting that confident phrasing may enhance perceived credibility even when content is inaccurate.

By combining behavioural adoption measures with eye-tracking data, this study provides new insights into how users evaluate AI-generated answers in search environments. The findings suggest that trust in AI-generated content is shaped less by surface-level credibility cues, such as logos or citation formatting, and more by the perceived quality, clarity, and plausibility of the content itself. These results contribute to the growing field of AI-supported information behaviour and underline the importance of designing search engine interfaces that prioritise factual transparency and content reliability over visual trust signals.

2. Acknowledgements

This section provides important disclosures and acknowledgements regarding the study 'Visual Cues and Trust in AI-Generated Search Results: An Eye-Tracking Study on Google AI Overviews', which was conducted at Hamburg University of Applied Sciences by Marius Brüning Genannt Wolter, Lisa Eleonora Haase, Benno Michaelsen, and Fabian Schott, and subsequently revised and edited by Dirk Lewandowski. The authors would like to express their gratitude to Sebastian Schultheiss for his valuable support in setting up the eye-tracking study and for serving as a reliable point of contact whenever technical software issues occurred.

Artificial intelligence tools were employed to support language editing, grammar checking, and text summarization during manuscript preparation. All research design, data analysis, and interpretations were conducted solely by the authors, who have reviewed and approved all AI-assisted text to ensure absolute accuracy and accountability.

Finally, this study did not receive any external funding, and no financial interests influenced its design, conduct, or reporting. The research was carried out at Hamburg University of Applied Sciences, a publicly funded institution, and no payments were made to the authors. The costs for participant incentives (Amazon vouchers) were covered internally by the university.

3. References

Amer, E., & Elboghdadly, T. (2024). The End of the Search Engine Era and the Rise of Generative AI: A Paradigm Shift in Information Retrieval. 2024 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), 374–379. https://doi.org/10.1109/MIUCC62295.2024.10783559

Ding, Y., Facciani, M., Poudel, A., Joyce, E., Aguinaga, S., Veeramani, B., Bhattacharya, S., & Weninger, T. (2025). Citations and Trust in LLM Generated Responses (Version 1). arXiv.https://doi.org/10.48550/ARXIV.2501.01303

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 2,500 participants. Proceedings of the 2003 Conference on Designing for User Experiences, 1–15. https://doi.org/10.1145/997078.997097

Kang, B., Höllerer, T., & O'Donovan, J. (2015). Believe it or Not? Analyzing Information Credibility in Microblogs. Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015, 611–616. https://doi.org/10.1145/2808797.2809379

Lewandowski, D., & Kammerer, Y. (2021). Factors influencing viewing behaviour on search engine results pages: A review of eye-tracking research. Behaviour & Information Technology, 40(14), 1485–1515. https://doi.org/10.1080/0144929X.2020.1761450

Liu, N. F., Zhang, T., & Liang, P. (2023). Evaluating Verifiability in Generative Search Engines (arXiv:2304.09848). arXiv. https://doi.org/10.48550/arXiv.2304.09848

Memon, S. A., & West, J. D. (2024). Search Engines Post-ChatGPT: How Generative Artificial Intelligence Could Make Search Less Reliable (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2402.11707

Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. Journal of Pragmatics, 59, 210–220. https://doi.org/10.1016/j.pragma.2013.07.012

Newman, N., Fletcher, R., Robertson, C., Ross Arguedas, A., & Kleis Nielsen, R. (2025). Reuters Institute Digital News Report 2025.

Ries, T. E., Bersoff, D. M., Baer, D., & Peterson, T. (2025). 2025 Edelman Trust Barometer Global Report.

Strzelecki, A. (2020). Eye-Tracking Studies of Web Search Engines: A Systematic Literature Review. Information, 11(6). https://doi.org/10.3390/info11060300

Trippas, J. R., Spina, D., & Scholer, F. (2025). Adapting Generative Information Retrieval Systems to Users, Tasks, and Scenarios. In R. W. White & C. Shah (Eds), Information Access in the Era of Generative AI (Vol. 51, pp. 73–109). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-73147-1_4