

Synchronous and Sequential Modes of SERP-based Relevance Evaluation

Jacqueline Sachse^a

^a *Humboldt-Universität zu Berlin, Unter den Linden 6, Berlin, 10099, Germany*

1. Introduction

Literature search constitutes an epistemic practice in which researchers assess the relevance of scientific work. While previous research has extensively studied the relevance clues guiding document selection (for an overview see Saracevic, 2016), less attention has been paid to how researchers sequence these clues during inspection of search engine results pages (SERPs). This study reconstructs researchers' SERP inspection to identify modes of relevance evaluation and examines how they vary across scientific disciplines.

2. Method

Life sciences and social sciences were selected as cases because both primarily engage in empirical research but differ in their epistemic cultures and patterns of scholarly communication (Becher & Trowler, 2001; Knorr-Cetina, 1999). Spanning various career levels, 25 researchers from each discipline participated in this study (life/social sciences: 3/5 professors, 17/14 postdocs, 5/6 predocs). A prefabricated SERP presented ten research articles with linked abstracts. Working on three scenarios requiring a literature overview in a given research area outside the participants' expertise, participants selected up to three documents for reading. Relying on think-aloud (Ericsson & Simon, 1993) and click data, relevance clues mentioned by participants were coded deductively drawing on Barry (1994), Barry and Schamber (1998), Behnert (2019) and Saracevic (2007).

3. Results

A code sequence analysis of relevance clues revealed two differing modes of SERP inspection: a 'synchronous' and a 'sequential' mode. Applying the synchronous mode, 28 participants (56%) performed one round of SERP inspection, commenting once on each rank while evaluating several relevance clues at a time (e.g., title, publication date, journal, and snippet simultaneously). Applying the 'sequential' mode, 22 participants (44%) focused only on a single relevance clue per rank in multiple rounds. For instance, screening each title in the first round, screening for high-impact journals in the second, for recent works in the third and for highly cited works in the fourth round of SERP inspection.

Moreover, Barnard's Unconditional Test revealed a statistically significant disciplinary pattern ($p=0.013$): While 72% of social scientists applied the synchronous mode, life scientists showed a more balanced distribution, with 60% applying the sequential mode. This suggests a more integrative approach to relevance assessment among social scientists, compared to a more criteria-specific filtering strategy among life scientists. Yet there was no significant variance in the frequency of document properties: All participants relied heavily on the title, followed by a wide margin by the journal and



publication date. Although Barnard's Unconditional Test indicated no statistical significance of *publication date* frequency between modes (sequential 13.04% vs. synchronous 10.33%, $p=0.08$), think-aloud data highlights qualitative differences between disciplines: life scientists more often aimed to identify recent works, whereas social scientists used it to locate a work within the overall discourse.

4. Discussion and Conclusion

Previous typologies of SERP inspection have focused on users' navigation through search results, for example by distinguishing breadth- and depth-first strategies (Barry & Lardner, 2011; Klöckner et al., 2004) or sporadic and structured navigational styles (Huang, 2020; Kinley et al., 2014). In contrast, the synchronous and sequential modes identified here describe procedural patterns of relevance evaluation, extending existing literature with a process-oriented perspective on SERP-based relevance evaluation. Quantitative and qualitative disciplinary comparison highlights literature search as an epistemic practice in which shared relevance clues are embedded in discipline-specific evaluative logics. Future research should examine whether the identified modes generalize across other disciplines and search environments.

5. Acknowledgements

This research has been funded under the research group grant "Reflexive Metrics" (FKZ 01PQ17002) by the German Federal Ministry of Education and Research (BMBF) as well as under the project "MultiSocVal: Wider societal value of research and consequences of its assessment: A multi-country and multi-method study" (funding reference no. 9C738) by the Volkswagen Foundation.

6. References

Barry, C. L. (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science*, 45(3), 149–159. [https://doi.org/10.1002/\(SICI\)1097-4571\(199404\)45:3<149::AID-ASI5>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1097-4571(199404)45:3<149::AID-ASI5>3.0.CO;2-J)

Barry, C. L., & Schamber, L. (1998). Users' criteria for relevance evaluation: A cross-situational comparison. *Information Processing & Management*, 34(2–3), 219–236. [https://doi.org/10.1016/S0306-4573\(97\)00078-2](https://doi.org/10.1016/S0306-4573(97)00078-2)

Barry, C., & Lardner, M. (2011). A Study of First Click Behaviour and User Interaction on the Google SERP. In J. Pokorny, V. Repa, K. Richta, W. Wojtkowski, H. Linger, C. Barry, & M. Lang (Eds.), *Information Systems Development* (pp. 89–99). Springer New York. https://doi.org/10.1007/978-1-4419-9790-6_7

Becher, T., & Trowler, P. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines* (2nd ed). Open University Press.

Behnert, C. (2019). Kriterien und Einflussfaktoren bei der Relevanzbewertung von Surrogaten in akademischen Informationssystemen. *Information - Wissenschaft & Praxis*, 70(1), 24–32. <https://doi.org/10.1515/iwp-2019-0002>

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data* (2nd ed.). The MIT Press. <https://doi.org/10.7551/mitpress/5657.001.0001>

Huang, Y.-H. (2020). Exploring Students' Search Behavior and the Effect of Epistemological Beliefs on Contradictory Issues. In F. F.-H. Nah & K. Siau (Eds.), *HCI in Business, Government and*



Organizations, 12204 (pp. 59–68). Springer International Publishing. https://doi.org/10.1007/978-3-030-50341-3_5

Kinley, K., Tjondronegoro, D., Partridge, H., & Edwards, S. (2014). Modeling users' web search behavior and their cognitive styles. *Journal of the Association for Information Science and Technology*, 65(6), 1107–1123. <https://doi.org/10.1002/asi.23053>

Klöckner, K., Wirschum, N., & Jameson, A. (2004). Depth- and breadth-first processing of search result lists. *CHI '04 Extended Abstracts on Human Factors in Computing Systems, CHI EA '04*, 1539. <https://doi.org/10.1145/985921.986115>

Knorr-Cetina, K. D. (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press. <https://doi.org/10.2307/j.ctvxw3q7f>

Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance. *Journal of the American Society for Information Science and Technology*, 58(13), 2126–2144. <https://doi.org/10.1002/asi.20681>

Saracevic, T. (2016). *The Notion of Relevance in Information Science: Everybody knows what relevance is. But, what is it really?* Springer. <https://doi.org/10.1007/978-3-031-02302-6>