A Triangulation Perspective for Search as Learning

Nilavra Bhattacharya[1], Jacek Gwizdka[1]

[1]School of Information, The University of Texas at Austin, USA

Abstract

Search engines and information retrieval (IR) systems are becoming increasingly important as educational platforms to foster learning. Modern search systems still have room to improve in this regard. We posit that learning-during-search is a good candidate for a human-centred metric of IR evaluation. This involves measuring two phenomena: learning, and searching. We discuss ways to measure learning, and propose a conceptual framework for describing searchers' knowledge-change during search. We stress the need for developing better measures for the search process, and discuss why we need to rethink the existing models of information seeking.

1. Introduction

As early as in 1980, Bertam Brookes, in his 'fundamental equation' of information and knowledge: $K[S] + \Delta I = K[S + \Delta S]$ had stated that a searcher's current state of knowledge, $K[S]$, is changed to the new knowledge structure, $K[S + \Delta S]$, by exposure to information ΔI, with the ΔS indicating the effect of the change [1, p. 131]. This indicates that searchers acquire new knowledge in the search process, and the same information ΔI may have different effects on different searchers' knowledge states. Fifteen years later, Marchionini described information seeking as "a process, in which humans purposefully engage in order to change their state of knowledge" [2]. Thus, we have known for quite a while that search is driven by higher-level human needs, and Information Retrieval (IR) is a means to an end, and not the end in itself.

When we consider information seeking as a process that changes the searcher's knowledge-state, the question arises whether the assessment of knowledge-acquisition-during-search, or *learning*, should subsume the standard IR evaluation metrics and the search interface usability metrics. It seems that to diagnose a problem or to understand a success of a search system, we would still need to control the standard aspects of a search system (e.g., results ranking, search user interface design features). However, a direct assessment of these "lower-level" aspects would lose on importance. On the other hand, support for more rapid learning across a number of searchers, and over a range of different search tasks can be indicative of an IR system that is more effective at supporting intelligence amplification and knowledge building [3]. In the last decade, this recognition that IR systems of tomorrow can become "rich learning spaces" and foster knowledge gain, has led to the emergence of the Search as Learning (SAL) research community [4], and the need to consider learning-during-search as a metric for evaluation of Interactive IR (IIR) systems.

2. Metrics for Learning & Knowledge

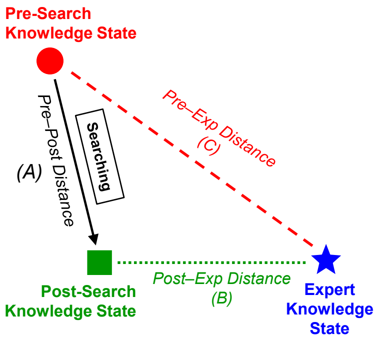
2.1. Experts vs. Novices

If we consider learning-during-search to be a good candidate for IR evaluation criterion, the next challenge is **how to measure learning**, or knowledge acquisition, possibly in an automated fashion. We can turn to educational psychology literature. A research report by the US National Research Council [5] identified the following key principles about experts' knowledge, illustrating the results of successful knowledge acquisition:

1. "Experts notice features and meaningful patterns of information that are not noticed by novices."
2. "Experts have acquired a great deal of content knowledge that is organized in ways that reflect a deep understanding of their subject matter."
3. "Experts' knowledge cannot be reduced to sets of isolated facts or propositions but, instead, reflects contexts of applicability: that is, the knowledge is 'conditionalized' on a set of circumstances."
4. "Experts are able to flexibly retrieve important aspects of their knowledge with little attentional effort."

Some of the above findings have been used by our community in the past. E.g, user learning has been measured by user's familiarity with concepts and relationships between concepts [6], gains in user's understanding of the topic structure [7], and user's ability to formulate more

Proceedings of the CIKM 2021 Workshops, November 1–5, Gold Coast, Queensland, Australia
nilavra@ieee.org (N. Bhattacharya); iwilds2020@gwizdka.com (J. Gwizdka)
https://nilavra.in (N. Bhattacharya); http://gwizdka.com (J. Gwizdka)
0000-0001-7864-7726 (N. Bhattacharya); 0000-0003-2273-3996 (J. Gwizdka)



Row No.	Pre-Post Dist (A)	Post-Exp Dist (B)	Pre-Exp Dist (C)	[Level of Knowledge *Change*] Intuition
1	Low	Low	Low	[Low] Expert
2	Low	Low	HIGH	X
3	Low	HIGH	Low	X
4	Low	HIGH	HIGH	[Low] Slow learner
5	HIGH	Low	Low	X
6	**HIGH**	**Low**	**HIGH**	[High] User gained new knowledge; **desirable**
7	HIGH	HIGH	Low	[High] **Knowledge loss:** started near expert-knowledge state before search; went further away from expert knowledge after search
8	HIGH	HIGH	HIGH	[High] **Mis-directed search/learning:** started away from expert-knowledge; went further away

Figure 1: A triangulation perspective for knowledge-change during search. We assume Pre-Search Knowledge, Post-Search Knowledge, and the reference Expert Knowledge as three vertices of a triangle (left figure). If we can compute the distance between the triangle's vertices, and then further dichotomize these distances as HIGH vs. Low, then we can have eight possible outcomes (right table). 'X' denotes outcomes violating the triangle inequality.

effective queries [8, 6]. From the above findings, we can think about ways to consider *Expert's Knowledge on the search topic* as 'gold-standard' or 'ground-truth' (by algorithmic parlance), for developing learning based IIR evaluation metrics.

2.2. Measuring Knowledge-Change

Recent literature on Search-as-Learning adopts three broad approaches to measure learning, or knowledge-change, with their own strengths and limitations. The **first approach** asks searchers to rate their self-perceived pre-search and post-search knowledge levels [9, 10]. This approach is the easiest to construct, and can be generalised over any search topic. However, self-perceptions may not objectively represent true learning. The **second approach** tests searchers' knowledge using factual multiple choice questions (MCQs). The answer options can be a mixture of fact-based responses (*TRUE*, *FALSE*, or *I DON'T KNOW*), [11, 12] or recall-based responses (*I remember / don't remember seeing this information*) [13, 14]. Constructing topic-dependant MCQs may take time and effort, which may be aided by automated question generation techniques [15]. For evaluation, this approach is the easiest, and often automated. However, MCQs allow respondents to answer correctly by guesswork. The **third approach** lets searchers write natural language summaries or short answers, before and after the search [16, 10]. Depending on experimental design, prompts for writing such responses can be generic (least effort) [17] or topic-specific (some effort) [15]. While this approach provides rich information about a searcher's knowledge state, evaluating such responses is the most challenging.

2.3. A Triangulation Perspective

We can conceptualize a triangle-based framework for searchers' knowledge-change during search (**Fig. 1**) Searchers initiate a search session with a Pre-Search Knowledge state. During search, they undergo a change in knowledge. On conclusion of search, searchers attain the Post-Search Knowledge state. We can attempt to measure this dynamic knowledge-change from a stationary reference point: Expert Knowledge on the search topic (ground-truth). If we imagine these three knowledge-states to be the three vertices of a triangle (Fig. 1, left), and if, by some hypothetical metric, we can compute the *distance* between any two of these knowledge-state points, then we have found a way to quantify learning-during-search.

Moving further, if we dichotomize the learning-during-search as 'HIGH' vs 'Low', by establishing a threshold value for the distances, then we can obtain eight possible knowledge-change situations (Fig. 1, right table). Three of these eight situations violate the triangle inequality[1] (denoted by 'X' in the table), and are therefore discarded. The remaining five valid situations are discussed below.

When Pre-Search Knowledge State and Post-Search Knowledge State are both very 'close' to Expert Knowledge (row 1 in table), we can assume the searcher is an **expert**. On the other hand, if Pre-Search Knowledge State and Post-Search Knowledge State are close to each other, but are far away from Expert Knowledge (row 4), the searcher is probably a **novice**, and also a **slow learner**, because on conclusion of search, their knowledge still remained far away from Expert Level. When

[1] sum of lengths of any two sides of a triangle is greater than the third side

the Post-Search Knowledge is closer to Expert than Pre-Search Knowledge (row 6), it implies that the searcher gained 'good amount' of new knowledge, and is thus, the **most desirable situation for Search as Learning**.

The last two rows of the table in Fig. 1 present two interesting, albeit undesirable, possibilities. If the Pre-Search Knowledge is closer to Expert, but the Post-Search Knowledge is further away (row 7), it can signify **knowledge loss** (which is also a form of knowledge *change*). On the other hand, if both the Pre-Search and the Post-Search knowledge are far away from Expert, and they are also far away from each other (row 8), then it is a case of **misdirected search**, and therefore, **misdirected learning**. A classic illustration of these two situations is health information seeking. Suppose a user is searching for cause and treatment of a small brownish spot on the wrist. If a physician examined the spot, they would immediately identify the spot to be caused by oil-splatter burn during cooking (Expert Knowledge State). The searcher may however, based on search results, come to the incorrect conclusion that they have skin cancer [18, 19]. Before the search, if the searcher correctly guessed that the spot was due to oil splatter burn, then the situation would be described by row 7 (knowledge loss, or increase in confusion), whereas if the searcher had no intuition about the cause of the spot before the search, the situation would be described by row 8. Both situations should be avoided by modern IIR systems.

2.4. Graph-based Operationalization

While the framework discussed in Section 2.3 is purely conceptual, we can think of a possible operationalization using graph-based representations, such as concept maps [20] or personalized knowledge graphs [21] (the terms are used interchangeably in this section).

"Learning does not happen all at once …it builds on and is shaped by what people already know" [3]. The Learning and the Cognitive Sciences have generally discovered that meaningful "deep learning" (of the human kind) requires learners to: (i) *relate* new ideas and concepts to previous knowledge and experiences; (ii) *integrate* knowledge into interrelated conceptual systems; and (iii) look for *patterns* and underlying principles [22, 23]. Concept maps are arguably, therefore, extremely suited to represent such knowledge structures, connecions, and patterns. A concept-map is a two-dimensional, hierarchical node-link diagram (graph) that depicts the structure of knowledge within a discipline, as viewed by a student, an instructor, or an expert in a field or sub-field. The map is composed of concept labels, each enclosed in a box (graph nodes); a series of labelled linking lines (labelled edges); and an inclusive, general-to-specific organization [24]. Concept-maps assess how well students see the "big picture", and where there are knowledge-gaps and misconceptions. They have been used for over 50 years to provide a useful and visually appealing way of illustrating and assessing learners' conceptual knowledge [25, 20, 24, 26, 27, 28, 29].

Expert knowledge or "ground truth" can be represented as topical knowledge-graphs of the information contained in online encyclopedias and knowledge bases. Searcher's pre and post-search knowledge states can be represented as concept maps or personal knowledge graphs. The searcher's graphs will evolve cumulatively over time, as the they encounter more information online. Construction of the personal knowledge graph can be manual (most effort), fully automated (least effort, but prone to prediction errors), or a human-in-the loop solution (an auto generated map is shown, but the user is free to modify it as necessary).

Having represented knowledge states as graph-based structures, measuring the similarity or distances between them becomes equivalent to the graph matching problem. Various algorithms and metrics have been proposed for exact and inexact graph matching [30]. Many of the solutions take an optimization-problem approach [31]. Some examples include structural similarity matching (comparing diameter, edges, distribution degrees etc.), iterative matching (comparing the node neighbours), subgraph comparison, and graph isomorphism [32].

Besides comparing two graphs, other kinds of analyses can reveal interesting patterns of learning and thinking, which can be correlated with search process measures. Some of these measures that have been used by Halttunen and Jarvelin [24] are addition, deletion, and differences in top-level concept-nodes, depths of hierarchy, and number of concepts that were ignored or changed fundamentally. In this regard, Novak and Gowin [25] have presented well-established scoring scheme to evaluate concept-maps: 1 point is awarded for each correct relationship (i.e. concept–concept linkage); 5 points for each valid level of hierarchy; 10 points for each valid and significant cross-link; and 1 point for each example. Such analyses methods can further inform the development of future operationalizations.

As our anonymous reviewers mentioned, knowing the goal of the learner is important in this scenario, as that will guide the formation of the learner's personal map. Furthermore, a search systems (or internet browsers) may provide a special 'learning mode' which is dedicated for measuring learning. This will help to avoid transactional or navigational search sessions that not necessarily aimed at learning/knowledge acquisition.

3. Measuring the Search Process

Learning-during-search involves two intertwined activities: learning, and searching. In Sec. 2, we discussed approaches to measure learning. The other part of the picture involves measuring the search process itself. Past research efforts has largely been devoted to measuring **search outcomes**: e.g., if a target document was reached, or if relevant results were shown. We argue that a more human-centred approach for measuring search is trying to quantify the **search process**.

3.1. Need for Longitudinal Studies

A major limitation of most IIR research efforts is that the user is examined in the short-term, typically over the course of a single lab session. The trend is similar in other HCI research venues. [33] stressed the need for longitudinal designs over a decade ago, yet a meta-analysis of 1014 user studies reported in the ACM CHI 2020 conference revealed that more than 85% of the studies observed participants for a day or less. To this day, "longitudinal studies are the exception rather than the norm" [34]. On the other hand, it is quite evident that knowledge acquisition is a longitudinal process, occurring gradually over time [3, 23, 5, 22]. Therefore, most educational curricula in schools and universities are spread across several months and years. "An over-reliance on short studies risks inaccurate findings, potentially resulting in prematurely embracing or disregarding new concepts" [34].

3.2. Need for Updated Theoretical Models

The Information Seeking literature is dominated by a large number of "multiple arrow-and-box" theoretical models. These models divide the information seeking process for complex search-tasks into different stages. Some argue that these models are not not "real models" but more of "short-hand common-sense task flows" [35, 36]. The mantra of these models have always been the same: they have "implications for systems design and practice". Unfortunately, these models, along with a significant body of IIR research, has not been able to go beyond suggestions, to providing concrete design solutions [37]. Moreover, there is great overlap in basic search strategies across many of these models [38], calling into question whether so many models are still relevant. Consequently, current search systems still predominantly use a "one-size-fits-all" approach: one interface is used for all stages of a search, even for complex search endeavours [39].

Again reiterating [33], we posit that these models, theorised decades ago for bulky desktop computers, are in need of improvement. Information seeking models have to incorporate the continuous or lifelong nature of online information searching, enabled by the proliferation of internet access in various handheld and portable digital devices. For instance, Marchionini's well known information seeking process (ISP) [2] models the information seeking behaviour into eight stages, with connecting feed-forward and feed-back loops between the stages. However, some researchers argue that users never really go "back" to an earlier state; e.g., "when reformulating the query, users do not really go back to the initial situation, they submit an improved query" [40]. With progress of time, there is continuous update of users' *information need* [41] and *search context* [42]. Thus, the intricate relationships between users' knowledge state, cognitive state, and other factors influencing search (search *context*), are ever-changing. Perhaps then Spink's model of the IR interaction process [43], which models interactive search as an infinite continuous process of sequential steps, or cycles[2], is better suited to explain information searching behaviour. Like time, there may not be an absolute beginning or end of a user's information searching process, but only search sessions. The user's cognitive state is always ever-changing and advancing, both during and between these search sessions. So a more realistic model will probably mean a fusion of Marchionini's and Spink's models, where Marchionini's entire ISP process becomes a *cycle* inside the Spink's model, with forward-directed arrows only. These types of realistic models, improved and validated by empirical data, will help to explain phenomena behind next-generation search interactions, such as searching and multi-tasking, multi-tabbed browsing [3, p. 36], multi-device searching, and multi-session searching [3, p. 61].

3.3. Neuro-physiological methods

Neuro-physiological methods (NP methods) [44] provide an interesting avenue to observe users while they interact with information systems. Two popular NP methods are eye-tracking [45, 46] and EEG [47]. Eye-tracking can capture eye-movements of users while they examine information on a screen. EEG captures (changes in) activation in different brain regions as users consume information. NP methods provide opportunities to understand and investigate how users gain knowledge during search. E.g, searchers use words or phrases they read in previous search results, in their future query reformulations [48]. Eye-tracking can detect and model this phenomenon. As a result, a number of recent efforts have tried to investigate learning (during search) using one or more NP methods [16, 17, 49, 50, 51, 15]. However, a major limitation of NP methods is that they (still) require lab environments for data collection. Taking lessons from the COVID-19 pandemic, as well as for scalability reasons, the IIR community needs search process metrics that can

[2] where each cycle consists of one or more interactive feedback occurrences of user's query input, IR system output, and user's interpretation and judgement of the output

measure remote user interaction, preferably over the long term. Consumer wearable devices (e.g., smartwatches) are a promising direction, since they can record physiological data such as heart rate, skin temperature, and galvanic skin response. White et al. [52] collected such data at a population scale, and correlated them with the population's search activities, to obtain improvements in relevance of result rankings.

4. Conclusion

The perspectives and propositions in this paper have been shaped by our experience in IIR research. The Information Processing Model from Educational Psychology states that information is most likely to be retained by a learner if it makes *sense*, and has *meaning* [53, p. 55]. When a piece of information *fits into the world-view* of the learner, it is said to make sense; when information is *relevant* to the learner, it has meaning. Our past research have primarily been in the second aspect of information retention: *relevance judgement.* After several user-studies and analysing multimodal sources of data, we generally conclude that relevant information attracts more visual attention, longer eye-dwell time, and more brain activations [45, 54, 17], compared to irrelevant information. Metrics which can capture the entire duration of an experimental trial, or the real-time flow of interactions, usually perform better as predictors, than metrics which aggregate the entire trial into a set of single numbers [46, 55, 54]. Hence we call for new and improved measures of the search process.

In the domain of Search as Learning, we employed word [56] and sentence [17] embeddings to semantically compare searcher's responses to expert knowledge. Word embeddings provided better visualization of results, showing clear separation of Pre-Search Knowledge from Post-Search and Expert Knowledge [56]. We also co-related Knowledge Change measures with interaction and eye-tracking measures. We saw that people who learnt 'less' spent more reading effort on SERPs [17]. Conversely, people who learnt 'more' were doing less reading overall; but most of their reading was on content pages. These high learners used more specialized terms in their queries, and reported higher mental workload (NASA-TLX).

In conclusion, we reiterate that learning-during-search is a good candidate for evaluating IR systems. We need more research to uncover relationships between the users' search process and their learning outcomes. Process measures can shed light on the various subtle aspects of human behaviour. If we understand them well, we can teach people to be more successful in their information seeking efforts, and maximize their learning outcomes. We envision that in the future, searchers will be able to 'track' and measure their knowledge progress over time, in a manner similar to tracking weight, fitness and physical exercises.

Acknowledgments

We thank the anonymous reviewers for their very helpful and thought provoking suggestions and feedback.

References

[1] B. C. Brookes, The foundations of information science. part i. philosophical aspects, Journal of information science 2 (1980) 125–133.

[2] G. Marchionini, Information Seeking in Electronic Environments, Cambridge University Press, 1995.

[3] R. W. White, Interactions with Search Systems, Cambridge University Press, 2016. doi:10.1017/CBO9781139525305.

[4] S. Y. Rieh, Research area 1: Searching as learning, https://rieh.ischool.utexas.edu/research, 2020. [Online; accessed 2020-04-19].

[5] N. R. Council, How People Learn: Brain, Mind, Experience, and School: Expanded Edition, The National Academies Press, Washington, DC, 2000. doi:10.17226/9853.

[6] P. Pirolli, P. Schank, M. Hearst, C. Diehl, Scatter/gather browsing communicates the topic structure of a very large text collection, in: Conference on Human Factors in Computing Systems (CHI'96), 1996.

[7] P. Zhang, D. Soergel, Process patterns and conceptual changes in knowledge representations during information seeking and sensemaking: A qualitative user study, Journal of Information Science 42 (2016).

[8] Y. Chen, Y. Zhao, Z. Wang, Understanding online health information consumers' search as a learning process, Library Hi Tech (2020).

[9] S. Ghosh, M. Rath, C. Shah, Searching as learning: Exploring search behavior and learning outcomes in learning-related tasks, in: Conference on Human Information Interaction & Retrieval (CHIIR), 2018.

[10] H. L. O'Brien, A. Kampen, A. W. Cole, K. Brennan, The role of domain knowledge in search as learning, in: Conference on Human Information Interaction and Retrieval (CHIIR), 2020.

[11] L. Xu, X. Zhou, U. Gadiraju, How does team composition affect knowledge gain of users in collaborative web search?, in: Conference on Hypertext and Social Media (HT'20), 2020.

[12] U. Gadiraju, R. Yu, S. Dietze, P. Holtz, Analyzing knowledge gain of users in informational search

sessions on the web, in: Conference on Human Information Interaction & Retrieval (CHIIR), 2018.

[13] S. Kruikemeier, S. Lecheler, M. M. Boyer, Learning from news on different media platforms: An eye-tracking experiment, Political Communication 35 (2018) 75–96.

[14] N. Roy, F. Moraes, C. Hauff, Exploring users' learning gains within search sessions, in: Conference on Human Information Interaction and Retrieval (CHIIR), 2020.

[15] R. Syed, K. Collins-Thompson, P. N. Bennett, M. Teng, S. Williams, D. W. W. Tay, S. Iqbal, Improving learning outcomes with gaze tracking and automatic question generation, in: The Web Conference (WWW), 2020.

[16] N. Bhattacharya, J. Gwizdka, Relating eye-tracking measures with changes in knowledge on search tasks, in: Symposium on Eye Tracking Research & Applications (ETRA'18), 2018.

[17] N. Bhattacharya, J. Gwizdka, Measuring learning during search: Differences in interactions, eye-gaze, and semantic similarity to expert knowledge, in: CHIIR'19, 2019.

[18] D. M. Australia, Why googling symptoms leads to cancer diagnosis and worse for your health, https://www.dailymail.co.uk/news/article-6927695, 2019. (Accessed on 02/07/2021).

[19] H. Times, Stop it! google is not a real doctor and no you don't have cancer, https://www.hindustantimes.com/health/story-EjdPqGmAe2CugNyNdUClSI.html, 2015. (Accessed on 02/07/2021).

[20] J. D. Novak, Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations, 2nd ed ed., Routledge, New York, NY, 2010.

[21] K. Balog, P. Mirza, M. G. Skjæveland, Z. Wang, Workshop on personal knowledge graphs, co-located with the 3rd automatic knowledge base construction conference (akbc'21), https://pkgs.ws/, 2021. (Accessed on 2021-10-14).

[22] R. K. Sawyer, The Cambridge handbook of the learning sciences, Cambridge University Press, 2005.

[23] S. A. Ambrose, M. W. Bridges, M. DiPietro, M. C. Lovett, M. K. Norman, How Learning Works: Seven Research-Based Principles for Smart Teaching, John Wiley & Sons, 2010.

[24] K. Halttunen, K. Jarvelin, Assessing learning outcomes in two information retrieval learning environments, Information Processing & Management 41 (2005) 949–972. doi:10.1016/j.ipm.2004.02.004.

[25] J. D. Novak, D. B. Gowin, Learning How to Learn, Cambridge University Press, 1984. doi:10.1017/CBO9781139173469.

[26] Y. Egusa, H. Saito, M. Takaku, H. Terai, M. Miwa, N. Kando, Using a Concept Map to Evaluate Exploratory Search, in: Proceedings of the Third Symposium on Information Interaction in Context, IIiX '10, ACM, New York, NY, USA, 2010, pp. 175–184. doi:10.1145/1840784.1840810.

[27] Y. Egusa, M. Takaku, H. Saito, How Concept Maps Change if a User Does Search or Not?, in: Proceedings of the 5th Information Interaction in Context Symposium, IIiX '14, ACM, New York, NY, USA, 2014, pp. 68–75. doi:10.1145/2637002.2637012.

[28] Y. Egusa, M. Takaku, H. Saito, How to evaluate searching as learning, in: Searching as Learning Workshop (IIiX 2014 Workshop), Regensburg, Germany, 2014. URL: http://www.diigubc.ca/IIIXSAL/program.html.

[29] Y. Egusa, M. Takaku, H. Saito, Evaluating Complex Interactive Searches Using Concept Maps., in: SCST@ CHIIR, 2017, pp. 15–17.

[30] Wikipedia, Graph matching, https://en.wikipedia.org/wiki/Graph_matching, 2021. (Accessed on 2021-10-14).

[31] X. Bai, Graph-Based Methods in Computer Vision: Developments and Applications: Developments and Applications, IGI global, 2012.

[32] Wikipedia, Graph isomorphism, https://en.wikipedia.org/wiki/Graph_isomorphism, 2021. (Accessed on 2021-10-14).

[33] D. Kelly, S. Dumais, J. O. Pedersen, Evaluation challenges and directions for information-seeking support systems, IEEE Computer 42 (2009).

[34] L. Koeman, Hci/ux research: what methods do we use? – lisa koeman – blog, https://lisakoeman.nl/blog/hci-ux-research-what-methods-do-we-use/, 2020-06-18. (Accessed on 11/08/2020).

[35] A. Dillon, Personal communication, 2020. [dated 2020-04-29].

[36] A. Dillon, No more information seeking models please, https://adillon.ischool.utexas.edu/2006/10/03/no-more-information-seeking-models-please/, 2006. [Online; accessed 2020-05-28].

[37] T. Saracevic, Information science, Journal of the American Society for Information Science 50 (1999) 1051–1063. doi:10.1002/(SICI)1097-4571(1999)50:12<1051::AID-ASI2>3.0.CO;2-Z.

[38] O. Hoeber, D. Patel, D. Storie, A study of academic search scenarios and information seeking behaviour, in: Conference on Human Information Interaction and Retrieval (CHIIR), 2019, pp. 231–235.

[39] H. C. Huurdeman, J. Kamps, From multistage information-seeking models to multistage search systems, in: Proceedings of the 5th Information

Interaction in Context Symposium, 2014, pp. 145–154.
[40] V. T. Tran, N. Fuhr, Using eye-tracking with dynamic areas of interest for analyzing interactive information retrieval, in: Conference on Research and Development in Information Retrieval (SIGIR'12), 2012.
[41] X. Huang, D. Soergel, Relevance: An improved framework for explicating the notion, Journal of the American Society for Information Science and Technology 64 (2013) 18–35. doi:10.1002/asi.22811.
[42] T. Saracevic, The Notion of Relevance in Information Science: Everybody knows what relevance is. But, what is it really?, Synthesis Lectures on Information Concepts, Retrieval, and Services (2016).
[43] A. Spink, Study of interactive feedback during mediated information retrieval., Journal of the American Society for Information Science (1997).
[44] J. Gwizdka, Y. Moshfeghi, M. L. Wilson, Introduction to the special issue on neuro-information science, Journal of the Association for Information Science and Technology 70 (2019) 911–916.
[45] J. Gwizdka, Characterizing Relevance with Eye-tracking Measures, in: Proceedings of the 5th Information Interaction in Context Symposium, IIiX '14, ACM, New York, NY, USA, 2014, pp. 58–67. doi:10.1145/2637002.2637011, 00028.
[46] N. Bhattacharya, S. Rakshit, J. Gwizdka, P. Kogut, Relevance prediction from eye-movements using semi-interpretable convolutional neural networks, in: Conference on Human Information Interaction and Retrieval (CHIIR'20), 2020.
[47] Z. Pinkosova, W. J. McGeown, Y. Moshfeghi, The cortical activity of graded relevance, in: Conference on Research and Development in Information Retrieval (SIGIR), 2020, pp. 299–308.
[48] C. Eickhoff, S. Dungs, V. Tran, An eye-tracking study of query reformulation, in: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '15, Association for Computing Machinery, Santiago, Chile, 2015, pp. 13–22. doi:10.1145/2766462.2767703.
[49] S. Kruikemeier, S. Lecheler, M. M. Boyer, Learning From News on Different Media Platforms: An Eye-Tracking Experiment, Political Communication 35 (2018) 75–96.
[50] J. Mao, Y. Liu, N. Kando, M. Zhang, S. Ma, How does domain expertise affect users' search interaction and outcome in exploratory search?, ACM Transactions on Information Systems 36 (2018).
[51] M. S. Franklin, J. M. Broadway, M. D. Mrazek, J. Smallwood, J. W. Schooler, Window to the wandering mind: Pupillometry of spontaneous thought while reading, Quarterly Journal of Experimental Psychology 66 (2013) 2289–2294.
[52] R. White, R. Ma, Improving search engines via large-scale physiological sensing, in: Conference on Research and Development in Information Retrieval (SIGIR'17), 2017.
[53] D. A. Sousa, How the Brain Learns, Fifth Edition, Corwin Press, 2017.
[54] J. Gwizdka, R. Hosseini, M. Cole, S. Wang, Temporal dynamics of eye-tracking and EEG during reading and relevance decisions, Journal of the Association for Information Science and Technology (2017).
[55] N. Bhattacharya, S. Rakshit, J. Gwizdka, Towards real-time webpage relevance prediction using convex hull based eye-tracking features, in: Symposium on Eye Tracking Research & Applications (ETRA '20), 2020.
[56] N. Bhattacharya, J. Gwizdka, Visualizing and quantifying vocabulary learning during search, in: Proceedings of the CIKM 2020 Workshops, October 19-20, 2020, Galway, Ireland, 2020.