

USER PERCEIVED LEARNING FROM INTERACTIVE SEARCHING ON BIG MEDICAL LITERATURE DATA

XIANGMIN ZHANG

School of Information Sciences, Wayne State University
106 Kresge Library, Detroit, MI 48202, USA

ABSTRACT. As in other fields, search engines have been heavily used as an information accessing tool for massive amount of medical literature data. This research investigates the user's learning during interactive searching process with the PubMed data, to find out what search behaviors would be associated with the user's perceived learning, and whether or not the user's perceived learning could be reflected in the existing search performance measures, so that such measures could also be used for indicating learning during searching process. The research used a data set collected by a research project on searching, which involved 35 participants at a major US university. The results show that the number of documents saved is significantly correlated with perceived learning for all search topics. None of the classical search performance measures is correlated with perceived learning. However, for specific topics, one of the performance measures, Recall, is significantly correlated with perceived learning. The results and the implications of the findings are discussed.

1. Introduction. The massive amount of data accessible today has dramatically enabled various sections of our society to make right decisions. However, it also brings challenges in a situation when a decision, based on available data, needs to be made quickly in response to an urgent need. Such situations include medical or healthcare circumstances when medical professionals need to determine the disease a patient may have, so that the patient could be appropriately treated. In such a case, quickly learning from the retrieved medical literature would be critical to settle the case. This research investigates people's learning behavior when using search engines, the major tools to access the massive amount of web data, including medical literature.

Information has increased explosively on the Web that is open to the public and is universally accessible. The effort of using the huge amount of information and data to support learning has been made in the past a couple of decades. Learning objects [6] have been created specifically for supporting learning, and various digital libraries consolidate information resources on the Internet in supporting learning [16]. Many people now learn directly from internet sources for different learning objectives (Goldman, et al., 2012 [10]). Self-directed learning, informal learning, life-long learning, as well as formal education in higher education, can happen as elearning in an information technology rich environment, which is more convenient for and easier access to learners.

2010 *Mathematics Subject Classification.* Primary: 97R50; Secondary: 97R71.

Key words and phrases. Information interactions, learning by searching, searching as learning, perceived learning, searching behavior.

This research considers search systems as a universally accessible IT tool that can play an important role in searching massive data on the web to support learning. While it is not difficult to find literature on the use of technology in general to support learning (e.g., Farwick, Hester & Teale, 2002 [9]; MacGregor & Lou, 2005 [14]; Shih, Chuang & Hwang, 2010 [19]; Edelson, Gordin & Pea, 1999 [8]), research on learning while conducting interactive searches is much needed, especially empirical research that can provide evidences on learning when accessing huge amount of information. In this study, we investigate into this emerging and important aspect of the learning when using information searching systems to access massive amount of information. Today's web search systems provide fast access to the massive amount of information on the Web, which would be impossible if without them. Naturally, searching has become a common activity in both formal classroom and in other informal learning settings. The combination of searching and learning has created the phenomena that is referred to as learning by searching (Yin, et al., 2013 [24]) or searching as learning in recent years (Rieh, et al., 2016 [18]; Vakkari, 2016 [20]). For this research, specifically, we are interested in the relationships between online searching and learning: whether there would be any user search behaviors that could be indicative of learning, and whether or not learning during searching could be assessed by using the existing search performance measures. While learning is a task by people at a wide span of ages, the population that is considered in this study is adult learners.

2. Searching and learning. Searching and learning behaviors have traditionally been two separate areas of studying. To lay a foundation for the current research, some discussion on learning and searching, and the relationship between the two, is necessary.

2.1. Learning. “Learning” can be defined from different angles (Ormrod, 2011 [17]), and in different types (Bransford, Brown & Cocking [4]), but in its general sense, “Learning is acquiring new or modifying existing knowledge, behaviors, skills, values, or preferences and may involve synthesizing different types of information,” as Wikipedia states. Ambrose and et. al. [1] defines learning as “a process that leads to change, which occurs as a result of experience and increases the potential for improved performance and future learning.” This definition emphasizes three critical components for learning: “1. Learning is a process, 2. Learning involves change in knowledge, beliefs, behaviors, or attitudes; and 3. Learning is not something done to students, but rather something students themselves do.” (p.3).

All these definitions for learning have one thing in common: learning is or results in change in the learner's knowledge, behaviors, beliefs, values, etc., and this change happens through a process.

2.2. Searching is part of learning process. Searching is a process of information seeking, in particular, for digital information. Because search always starts with an information need or statement, which represents the user's intended information, searching also involves the evaluation or judgement on the search results to see if the search results are indeed related to the search objectives. A complete search process normally includes the following steps: forming the search query (or transforming the internal information need into a formal, explicit search statement); submitting the search query to the search system; evaluating the search results; if

not satisfied or the information need changed, revising the search query and resubmitting it, and reevaluating the results. This information searching process has long been recognized as part of the learning process because the information retrieved is used as the input for learning. Kuhlthau (2004) [13] develops an information search process model describing different stages during the search process that people seek information to deepen and broaden their understanding of the world around them. This model is part of people's learning process. If the search system could not provide the needed information in the learning process, problems may arise which would hinder learning.

Despite the recognition of the importance of searching or information seeking for learning, the traditional view about searching and learning is: they used to be two separate things: searching is to collect information for learning, and learning is another process that uses the information being collected.

2.3. Learning by searching. Searching is not just part of learning process. Searching has become more than just finding pieces of information: it shares learning activity features. People, particularly students, often employ explicit search as part of the learning process in studying of a specific topic. Marchionini (2006) [15] argues that learning is a key process within the common activity of exploratory search, among three kinds of search activities: lookup (finding a fact, etc.), learning (acquisition of knowledge, etc.), and investigation (analysis, evaluation, discovery, etc.). As more primary materials go online, searching to learn is increasingly viable. Exploratory search systems are needed to support the full range of users search activities, especially learning and investigation, and not just lookup, which usually can be completed by one iteration of query-results. This inseparability of learning and searching is also promoted by informed learning (Bruce & Hughes, 2010 [5]), in which it is argued that information activities and learning are simultaneous processes.

The notion of Anomalous State of Knowledge (ASK) (Belkin, Oddy & Brooks, 1982 [3]) explains the internal motivation for learning when people need to search for information. According to the theory of ASK, a user's information need that motivates the user to search, is a gap in knowledge that the user needs to fill up. Therefore, searching information is the process to bridge the knowledge gap, and is thus a learning process. People have an inexpressible need for information, which cannot be explicitly specified when they are in ASK, and can only be learned. On the other hand, the search system should present something new to the user so that the user could learn to fill in their knowledge gap.

Further evidence to support that searching is learning process is provided by Jansen, Booth & Smith (2009) [12], based on the cognitive processes involved. The cognitive processes involved in learning are summarized in Anderson & Krathwohl (2001) [2], and are arranged in the authors' "Taxonomy Table." Anderson & Krathwohl (2001) [2] lists six major categories of cognitive processes: remember, understanding, apply, analyze, evaluate and create, in the order of from simple to complex. Based on these cognitive processes, Jansen, Booth & Smith (2009) [12] classify 426 searching tasks according to cognitive process features identified by the Taxonomy Table in Anderson & Krathwohl (2001) [2]. Seventy-two participants were asked to perform these tasks in a laboratory experiment. The results show that information searching is a learning process with searching behavior characteristics specific to particular cognitive levels of learning. The results indicate that applying and analyzing, the middle two of the six categories, generally take the most

searching effort in terms of queries per session, topics searched per session, and total time searching. The lowest two learning cognitive processes, remembering and understanding, exhibit searching characteristics similar to the highest order learning cognitive processes of evaluating and creating. Based on the findings, Jansen, Booth & Smith (2009) [12] suggest that a learning theory may be better to describe the information searching process. Such a theory, however, will need to be based on more understanding of learning during the search process.

In the present article, “learning (by searching)” means acquiring new knowledge about a topic through interactive searching activities. “Learning by searching” reflects the emerging trend that learning, particularly informal learning, is increasingly blended in the interactive information searching process. This can be evidenced by the use of digital content as well as search engines in classrooms: students often employ searching as part of their classroom learning process when studying a specific topic. In this scenario, learning happens during the search process, not after searching.

There has been some research on learning by searching. To support learning by searching, Yin, et al. (2013) [24] designed a system that could do automatic analysis on the search results for a particular course. This system combined searching and learning automatically, and helped address the fundamental issue of supporting learning while searching.

Zhang, et al (2014) [25] investigate the factors (other than the behaviors) that may be associated with perceived learning. These factors include the user’s prior knowledge, prior search skills/experience (because the searches were on genomics documents, search skills/experience was restricted to Medline database search experience), search task characteristics such as general/specific, and the user satisfaction with the search results they found in relating to a specific search topic. The results show that in general (without considering the task characteristics), all three factors: prior knowledge, prior search experience, and satisfaction, are significantly correlated with perceived learning. But only satisfaction is found to be a significant factor contributing to perceived learning, by GLM/ANOVA analysis.

Closely related to investigating learning by searching, Goldman, et al. (2012) [10] used think-aloud method to study internet source evaluations and reading behaviors related to learning, on the observation that Readers increasingly attempt to understand and learn from information sources they find on the Internet. In the study, 10 better learners were contrasted with 11 poorer learners. Results indicate that better learners engaged in more sense-making, self-explanation, and comprehension-monitoring processes on reliable sites as compared with unreliable sites, and did so by a larger margin than did poorer learners. Better learners also engaged in more goal-directed navigation than poorer learners. This study, however, did not investigate the searching aspect. Walraven, Brand-Gruwel & Boshuizen, (2009) [22] and Willoughby, et al. (2009) [23] investigated evaluating and using information when searching on the internet. Learning could be implied in the study but it would be better to be explained explicitly.

3. Research questions. Based on the previous work, the current work seeks to extend the scope of previous research to the understanding of how users learn by searching. In addition, by investigating the relationships between users’ learning and their actual search performance in terms of search outcome, the current research intends to find out if the existing search performance measures could be used to

assess learning during searching. It should be noted that although searching and learning share common cognitive processes and are considered inseparable in this study, the learning tasks are normally not explicitly defined, but are implied in searching tasks or topics. Therefore, when investigating learning during searching process, instead of using some actual learning tasks, this study uses “perceived learning,” i.e., the user’s feelings about their knowledge gains from the searching process. Two research questions are addressed in this research:

1. What search behaviors are associated with perceived learning? Users behaviors on a search system have been studied for decades. These include querying behaviors, search result accessing behaviors, and so on. Behaviors which are significantly associated with perceived learning need to be identified. This identification will help understand when and how when users conduct interactive searching, they are actually learning.
2. Are user’s learning correlated with the search success as measured by typical search performance measures, so that learning during searching could also be indicated by using such measures?

For learning during interactive searching process, one possible solution for learning assessment would be using the existing measures for search performance to assess learning. Search performance has traditionally been evaluated against the number of relevant documents retrieved, using the classic measures of Precision, Recall, and F measure (these measures are explained later in the METHODS section). These “relevance” based measures test the user’s ability to find relevant documents using the search system. They do not relate to learning directly. However, it was unknown if these measures would be able to be used for assessing learning by searching. Duggan & Payne (2008) [7] demonstrated the correlation between users knowledge, represented by a pre-knowledge score, and search performance, represented by a search score. Participants took a knowledge test before searching and were scored by the proportion of items correctly answered. The same test was applied after searching. It was found that the score after searching was positively correlated with the pre-search knowledge score. Duggan & Payne (2008) [7], however, did not explicitly evaluate the difference between the search score and knowledge score. This study will compare these measures with the user’s perceived learning, hoping to find out if the measures could also be applied to assessing learning in searching.

By answering the above research questions, the current research seeks to further the understanding of learning by searching, and to provide evidence for developing needed search technologies to support learning.

4. Methods. This research uses the data collected and shared by a large research project on user’s information searching behavior at a major US research university. A detailed and complete description of the research design and the user experiment for the data collection can be found in (Zhang, et al. 2014) [25]. In this article, we describe the resulting data and the measures that are used in this study.

4.1. Data set. The data was collected through a laboratory user experiment in which 35 participants performed four search tasks using the standard Text Retrieval Conference (TREC) Genomics Track data, which are PubMed documents. The search topics were also in genomics and were adopted from TREC, Genomics Track dataset (Hersh & Voorhees, 2009) [11]. The participants’ search behaviors were recorded by the system. Before and after each searching task, participants were also asked to fill out the pre- and post-task questionnaires. The logged behavior

data and the completed questionnaires consist of the primary data used in this research.

Based on the original research design (Zhang, et al., 2014) [25], the collected data set includes the following types of data on two sets of search topics – general and specific topics:

- Users’ search behavior data, such as the number of queries submitted to the system for a search topic, the average query length for a topic, the time spent on a topic, the documents selected from the search results pages (SERPs), and so on. A complete list of search behavior variables used in this study is listed in Table 2.
- Users’ search performance data, mainly the number of documents judged and saved by the user as relevant for a topic, and
- The questionnaire data, from both the pre- and post-task questionnaires, which included questions on demographic data and a question of if the user felt that new knowledge was learned through the search on a topic.

Because the documents used in the research were in the genomics area and were part of the PubMed database, the National Library of Medicine’s controlled vocabulary, Medical Subject Headings (MeSH) tree, was used to determine the specificity of the search topics in the study. A search topic’s specificity was determined by the level of the topic subject in the MeSH tree, that is, the length of path to the root in the MeSH category tree. A topic was classified into the general category if the topic subject’s level in MeSH hierarchy was above 3. Below the first 3 levels, a topic was considered a specific one. Given the variability of hierarchical levels in different parts of MeSH, specificity could be further distinguished. However, the small number of topics in this study made it unnecessary to further distinguish levels below 3. The search topics used in the study are listed in Table 1 below.

TABLE 1. Search Topics and Their Specificity

TREC topic #	Topic title keywords	MeSH category	Specificity
2	Generating transgenic mice	Genetic structure	Specific (4)
7	DNA repair and oxidative stress	Genetic processes	General (1)
42	Genes altered by chromosome translocations	Genetic phenomena	Specific (4)
45	Mental Health Wellness-1	Genetic phenomena	General (1)
49	Glyphosate tolerance gene sequence	Genetic structure	General (1)

In Table 1, topics 2 and 42 are specific topics because their topic subjects are located at level 4 respectively in their correspondingly MeSH hierarchical tree (the number in the parentheses indicates the level in the subject hierarchy). Topics 7, 45 and 49 are general topics because they are all at the first level, the highest level of the corresponding part of the MeSH tree. Although there were 5 topics, each participant was presented and asked to do only 4 topics in the experiment. Topics 42 and 49 were substituted between subjects. A complete description of the search topics is attached in the Appendix.

4.2. **Measures.** The following measures are used in the study:

Perceived Learning

This is a user self-reported rating on a 7 point scale in the post-task questionnaire, to the question if new knowledge was learned from the search, from 1 for Not at all, 4 for Somewhat, and 7 for Learned a lot (Extremely).

Search Interaction behaviors/activities

Three categories of search behaviors are considered in this study: querying behaviors, such as the number of queries, average query length, and average query time; document viewing behaviors, such as the number of documents being saved, being viewed, the ranking position of the documents being opened, average document dwell time, etc.; and the general task interaction behaviors, such as the number of actions per task, task completion time, and so on. In total, 11 behavior variables were analyzed in this study, as listed in Table 2.

These behavior variables have been frequently used in information seeking research.

Search performance measures

Three classical performance measures are used in the study: Precision, Recall, and F measure (Van Rijsbergen, 1979) [21]. Precision p is the number of correct search results divided by the number of all retrieved results, and Recall r is the number of correct search results divided by the number of all possible relevant results in the search system.

TABLE 2. Behavior Variables

Behavior Variables	Description
# of Qs	The total number of queries submitted to the search system for a specific search task
q-Length	Query length is the number of words contained in a query. Here query length is the average length of multiple queries for a search task
# of Docs. saved	Number of documents/abstracts saved from the search results for a task
# of Docs. viewed	Number of documents/abstracts opened and viewed from the search results for a topic
Ratio-of-DocsSaved/Viewed	The ratio of documents saved and the documents opened/viewed
# of Actions task	The total number of actions during working on a search topic. The actions include both keyboard and mouse actions
# of SERPs viewed	Number of search result pages viewed or checked that were returned by the search system
Time for the Task	The total time spent on tasks
Ranking on SERPs	The average ranking position of the documents opened in SERPs. "1" is the top ranking, most related by the system and the larger the number, the lower the ranking is.
Average dwell time	Average time spent on viewing document/abstract
Querying time	Average time spent on working on queries

Each participant's performance measures are included in the dataset. These performance measures were calculated based on the participants' evaluation of search results. In the experiment that generated the data set, Participants were asked to conduct searches on the experimental system and to find and save as many relevant documents as possible. After finishing their search activities, they evaluated all saved documents. Participants rated the relevance of the saved documents using a five point scale ranging from "not relevant" to "highly relevant" with "somewhat relevant" as the mid-point.

As part of the data from TREC, which includes documents to be searched and the search topics, TREC also provides relevance judgements for the documents related to each search topic, on a 3 point scale: from "not relevant," "somewhat relevant," to "highly relevant." Using the TREC assessment format, the participant judgments were mapped to the TREC scale with ratings of 4 or 5 as "highly relevant", and ratings of 2 or 3 as "somewhat relevant."

Participant performance for each topic was calculated by checking for agreement with the TREC assessment. Both "somewhat relevant" and "highly relevant" were taken as relevant when calculating precision and recall. Recall was calculated by dividing the number of (correct) relevant documents by the number of TREC-assessed relevant documents for the task.

The F score is the weighted average of recall and precision, where an F1 score reaches its best value at 1 and worst score at 0. We use the F2 score, which weights recall twice as much as precision, because our tasks are recall-oriented: participants were asked to find all related articles.

It should be noted that the MeSH index terms in documents were not available to the user when they viewed the search results (abstracts). The relevance criteria for participants were presumably based on their interaction with the available text.

4.3. Data analysis. Pearson correlation analysis and GLM/ANOVA procedure are the main statistic methods used in the study. While correlations can find associations between two variables, there's no causal relationship that can be determined from the correlation analysis results. Therefore, GLM/ANOVA analysis is also applied. The data analyses were performed using SPSS.

5. Results and discussion.

5.1. Perceived learning and search behaviors. From the perspective of designing information systems to support learning, an important issue is to identify the behaviors that are associated with learning new knowledge through searching, so that such behaviors could indicate the users learning. The results of correlation analysis are presented in Table 3.

As presented in Table 3, among all the 11 behavior variables investigated, only three variables have significant correlations with perceived learning: the number of documents saved, the ratio of the number of saved documents and the number of documents opened or viewed, and the average ranking position of the documents opened.

TABLE 3. Correlations between Perceived Learning and Behavior Measures

Behavior Variables		Correlation with Perceived Learning ($n = 140$)
# of Qs	Correlation	-.085
	Sig. (2-tailed)	.320
q length	Correlation	.060
	Sig. (2-tailed)	.482
# of Docs Saved	Correlation	.180 ¹
	Sig. (2-tailed)	.034
# of Docs opened or viewed	Correlation	.082
	Sig. (2-tailed)	.336
Ratio of Docs Saved or viewed	Correlation	.311 ²
	Sig. (2-tailed)	.000
# of Actions task	Correlation	.107
	Sig. (2-tailed)	.206
# of SERPs viewed	Correlation	.086
	Sig. (2-tailed)	.314
Time for task	Correlation	.031
	Sig. (2-tailed)	.714
Ranking on SERPs	Correlation	.168 ¹
	Sig. (2-tailed)	.047
Average dwell time	Correlation	-.047
	Sig. (2-tailed)	.584
Query time	Correlation	-.049
	Sig. (2-tailed)	.567

¹ Correlation is significant at the 0.05 level (2-tailed).

² Correlation is significant at the 0.01 level (2-tailed).

Of the three variables, the ratio of the documents saved from all viewed is significant at $p < 0.01$ level. The other two are significant at the 0.05 level ($n = 140$). All correlations are positive. The results indicate that when more relevant documents are being saved, without viewing too many documents, it is more likely the user will report that they have learned. Since the number of documents opened alone does not have a significant correlation with perceived learning, it seems that the significance of the ratio mainly comes from the contribution of the number of documents saved. The significant correlation between the ranking position and perceived learning indicate that the lower (the larger the number) the mean rank of the opened documents in SERPs, the more likely the user will feel they have gained new knowledge.

A follow-up GLM/ANOVA analysis identified the ratio is the only variable that has a significant effect on perceived learning ($F = 10.838$, $p = .001$). All other behavior variables do not show significant effect on perceived learning. The test results are listed in Table 4.

The results show that stronger perceived learning is associated with more document savings and lower ranking in the SERP list, which imply that more effort is needed during the search interaction process. These two factors could be indicators of a user's learning. The finding sheds light on how the system may predict how much users gain knowledge through observable search behaviors.

No significant correlations are found between perceived learning and other user behaviors or efforts, such as the amount of time spent on the task, time spent on each page, number of pages viewed, number of queries issued, etc. Intuitively, learning might be associated with some of these behaviors. Future work will need to continue examining the relations between learning and these behaviors. If some additional important behaviors could be identified, it will help the system infer users learning, and adapt search accordingly for the user.

5.2. Perceived learning vs. performance. Interestingly, the correlation analysis found that perceived learning is not necessarily associated with any of the three search performance measures: Precision, Recall and F_2 (that is biased towards Recall). There are no significant correlations between the two sets of measures. The results are listed in Table 5. A follow-up GLM/ANOVA analysis found no significant effect on perceived learning from any of the performance measures.

TABLE 4. GLM/ANOVA Results of Effects of Behaviors on Perceived Learning

Dependent Variable	Perceived Learning				
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	44.296 ^a	11	4.027	2.102	.024
Intercept	.221	1	.221	.115	.735
# of Qs	3.299	1	3.299	1.722	.192
q length	.011	1	.011	.006	.939
# of Docs Saved	1.068	1	1.068	.558	.457
# of Docs opened viewed	1.967	1	1.967	1.027	.313
Ratio of Docs Saved Viewed	20.760	1	20.760	10.838	.001
#ofActions task	.309	1	.309	.161	.688
# of SERPs viewed	1.426	1	1.426	.745	.390
Time for Task	2.610	1	2.610	1.362	.245
Ranking on SERPs	.505	1	.505	.264	.608
Average dwell time	6.798	1	6.798	3.549	.062
Query time	7.339	1	7.339	3.831	.052
Error	245.175	128	1.915		
Total	2434.500	140			
Corrected Total	289.471	139			

^a R Squared = .153 (Adjusted R Squared = .080)

One possible explanation for the result could be that a document judged relevant to the search topic does not necessarily add new knowledge to the user, i.e., the relevant document is not connected to the users' ASK status. "Learning" new knowledge is a goal different from finding "related" documents. A document could be "relevant" in many ways.

This result has significant implications and highlights a gap in search performance evaluations. If learning is important for searching, the performance measures should take learning measures into account. The measures that focus only on search results may be unable to indicate the learning results. Some assessment mechanism is needed for search systems that could support for user learning during the search process.

TABLE 5. Correlations between Perceived Learning and Performance Measures

Performance Measures		Correlation with Perceived Learning (n=140)
Precision	Correlation	-.069
	Sig. (2-tailed)	.416
Recall	Correlation	.052
	Sig. (2-tailed)	.545
F2Score	Correlation	.047
	Sig. (2-tailed)	.579

5.3. Effects of topic characteristics on perceived learning. The analyses described in the above sections did not consider the search task characteristics, i.e., the difference between general topics and specific topics, which intuitively is related to learning. The data was further separated into two subsets, one for general topics and the other one for specific topics. The same statistical analyses were conducted separately for each of the two types. The purpose was to find out if this general-specific topic characteristic would be associated with perceived learning. The number of participants included in the data set for the general topics is different from that of specific topics, due to the unbalanced number of topics in each category. The correlations between perceived learning and search behaviors are presented below first, which is followed by the results of correlation analysis on search performance.

5.3.1. Search behaviors. The analysis on the whole data set found that three behavior variables had significant correlations with perceived learning, which were the number of documents saved, the ratio of the number of documents saved and the number of documents viewed, and the ranking position of the relevant/saved documents on SERPs. The results here are slightly different from the results in Section 5.1. Table 6 presents all correlations between behavior variables and perceived learning, under the two conditions separately: one for general topics and one for specific topics.

As Table 6 shows, for both general and specific topics, the average ranking on SERPs is no longer a significant factor that is associated with perceived learning. One direct consequence of slitting the data into two subsets is that the sample size in either subset is much smaller than that in the whole data set. It is possible that while checking down the search result list may be associated with perceived learning in large samples, it may not be the case for smaller samples. Given that one individual user's data size is normally small, a document's ranking position in SERPs may not be an important factor to consider.

Interestingly, the number of saved documents was found significantly correlated with perceived learning in the whole data set. But this further analysis found

that it actually is significant only for the specific topics, not for general topics. It could be the case that specific topics are easier to learn than general topics because specific topics are relatively clearer than general ones, which normally are vaguer than specific ones.

The ratio is still significantly correlated with both general and specific topics. However, different from the test results from the whole data set where a significant effect is found with the ratio on perceived learning, a GLM/ANOVA analysis does not found significant effect of this variable. Again, it may be because the samples are not big enough to draw the results. In fact, none of the behavior variables is found to have an effect on perceived learning in either the general or specific topic case.

TABLE 6. Correlations between Perceived Learning and Behavior measures for General and Specific Topics Separately

Behavior Variables		Correlation with Perceived Learning	
		General Topics (n=90)	Specific Topics (n=50)
# of Qs	Correlation	-.085	-.015
	Sig. (2-tailed)	.320	.915
q length	Correlation	.045	.084
	Sig. (2-tailed)	.671	.563
# of Docs Saved	Correlation	.126	.338
	Sig. (2-tailed)	.235	.016
# of Docs opened or viewed	Correlation	.045	.179
	Sig. (2-tailed)	.671	.213
Ratio of Docs Saved or viewed	Correlation	.248	.457
	Sig. (2-tailed)	.018	.001
# of Actions task	Correlation	.058	.0234
	Sig. (2-tailed)	.589	.102
# of SERPs viewed	Correlation	.037	.199
	Sig. (2-tailed)	.731	.165
Time for task	Correlation	-.014	.109
	Sig. (2-tailed)	.899	.453
Ranking on SERPs	Correlation	.192	.122
	Sig. (2-tailed)	.069	.399
Average dwell time	Correlation	-.017	-.105
	Sig. (2-tailed)	.871	.469
Query time	Correlation	-.092	.011
	Sig. (2-tailed)	.390	.942

5.3.2. *Performance measures.* A GLM/ANOVA analysis was first conducted to examine if the search topic characteristic would have significant effect on participants' perceived learning, if a given topic is general or specific. Similar to the result from the analysis on the whole data set in Section 5.2, no significant effect is found on perceived learning from performance measures in either general or specific topic case.

While the result is not significant from the GLM/ANOVA analysis, the correlation analysis found some meaningful results that show significant correlations with perceived learning. The results are presented in Table 7.

Table 7 shows that both Recall and F_2 scores have significantly positive correlations with perceived learning for specific tasks, at $p = 0.037$. Because the F_2 measure is biased toward Recall, as we introduced earlier in this article, the significant correlation is mainly with Recall, which for some reason the GLM/ANOVA was unable to detect. This finding is different from the result with the whole data set, where no significant correlations were found with any of the performance measures, despite a larger sample size of whole data set. It shows that when searching for specific topics, whether or not to find as many as possible all the relevant documents does seem to be related to whether or not the user would feel having learned new knowledge.

TABLE 7. Correlations between Perceived Learning and Different Types of Topics

Performance Measures	Perceived Learning	
	General (n=90)	Specific (n=50)
Precision	$r=-.083$, $p=.439$	$r=-.040$, $p=.785$
Recall	$r=.021$, $p=.842$	$r=.296$, $p=.037$
F_2 Score	$r=.015$, $p=.888$	$r=.295$, $p=.037$

Performance measures do not seem to be strongly correlated to perceived learning for search topics in general. Only for specific topics are there significantly positive correlations between Recall and perceived learning. F_2 scores, which are biased towards Recall, and perceived learning, are also significantly correlated.

It could be that in the case of specific topics users are able to gain more concrete ideas and, thus, are more capable of identifying the relevant documents, are thus able to learn. In general topics, users may not be able to learn much in abstract terms on the general topic. They might also have difficulty in gaining a clear idea of how much they had learned about a general topic than from working on a specific topic.

6. Conclusions. One of the goals for today's access to massive amount of data and information is to learn. Accomplishment of this goal is particularly important when a fast response is needed, such as in an urgent medical or healthcare situation. With today's huge amount of information that is universally accessible on the web through search engines, more research is needed to investigate how people access the information in enhancing their learning. This research investigated the user's learning issues when interacting with digital content through interactive searching, which is an important use of accessible information technology in today's world. In response to the two research questions, the research found that:

- Perceived learning is only associated with limited types of search behaviors: the number of documents saved (as relevant). The more the saved documents, the strong feeling of having learned. The ranking position of the documents opened in SERPs can also significantly correlate with perceived learning, but only if the sample size is large: the lower ranking positions of the documents

opened, the more the user would perceive learning. But when the sample size is smaller, the correlation is not strong. Realistically for an individual user, this perhaps means the ranking position is not a strong behavior factor to consider as an indicator of learning.

- Perceived learning does not show, in general, significant correlation with search performance, measured by the classical information retrieval metric: Precision, Recall and F_2 measures. However, considering the search topic characteristics and focusing on specific topics, Recall is significantly correlated, at $p = 0.05$ level, with perceived learning. This result suggests that for a specific search topic, a user's learning is related to the number of relevant documents the user can find: the more relevant documents found the more the user may learn. Recall may be used as an indicator of the user's learning, if the search topic's specificity could be determined.

It should be noted that since most of the statistical significance were found from the correlation analyses, not from GLM/ANOVA analyses, the relationships between the significantly correlated variables and perceived learning may not be causal ones. In fact, given the complexity of both learning and searching processes, with many cognitive processes involved, it probably is premature trying to find causal relationships between the two.

It should also be admitted that the study focuses on a narrow domain: genomics. Therefore, the findings may not be appropriate to generalize to other subject areas. Similar research is needed in other areas to collect empirical evidences.

As the amount of accessible data/information has been steadily increasing in the past a few decades, studying people learning that occurs when they use information access tools such as search systems is increasingly important. The results of the study advance the state of knowledge in information accessing-related fields, and should also be interesting to evaluating the impact of universal access to information in society. Practically, the results of the study also have significant implications for search-based technological support for accessing information for learning: such tools may need to incorporate new learning supportive functions. While artificial intelligence techniques may solve some issues, but at the current stage of computing power, supporting human learning from big data is still a great challenge for the design of many information systems. It is hoped that, in today's big data era, an information access tool could not only fast return a list of hits that are based on the algorithms, but also could support human learning.

Acknowledgment. This article is an extended version of the author's paper presented at the HCI International Conference 2016.

The author thanks all other members of the Personalization of the Digital Library Experience (PoODLE) research team at Rutgers University for collecting and sharing the data, without whose efforts this work could not have been accomplished.

REFERENCES

- [1] S. A. Ambrose, M. W. Bridges, M. DiPietro, M. C. Lovett and M. K. Norman, *How Learning Works: Seven Research-Based Principles for Smart Teaching*, Jossey-Bass, A Wiley Imprint. P. 3, 2010.
- [2] L. W. Anderson and D. A. Krathwohl, *A Taxonomy for Learning, Teaching, and assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, New York: Longman, 2001.
- [3] N. J. Belkin, R. N. Oddy and H. M. Brooks, [ASK for information retrieval: Part 1: Background and theory](#), *Journal of Documentation*, **38** (1982), 61–71.

- [4] Bransford, J. D., Brown, A. L. and R. R. Cocking, *How People Learn: Brain, Mind, Experience, and School*, Washington: National Academies Press, 2000.
- [5] C. Bruce and H. Hughes, [Informed learning: A pedagogical construct attending simultaneously to information, use and learning](#), *Library & Information Science Research*, **32** (2010), A2–A8.
- [6] Downes, *Learning Objects: Resources For Distance Education Worldwide*, International Review of Research in Open and Distance Learning, 2001.
- [7] G. B. Duggan and S. J. Payne, [Knowledge in the head and on the web: Using topic expertise to aid search](#), *CHI '08 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2008, 39–48.
- [8] D. C. Edelson, D. N. Gordin and R. D. Pea, Addressing the challenges of inquiry-based learning through technology and curriculum design, *Journal of the Learning Sciences*, **8** (1999), 391–450.
- [9] R. Farwick, O. J. L. Hester and W. H. Teale, Where do you want to go today? Inquiry-based learning and technology integration, *The Reading Teacher*, **55** (2002), 616–625.
- [10] S. R. Goldman, J. Braasch, J. Wiley, A. Graesser and K. Brodowinska, Comprehending and learning from internet sources: Processing patterns of better and poorer learners, *Reading Research Quarterly*, **47** (2012), 356–381.
- [11] W. Hersh and E. Voorhees, [TREC genomics special issue overview](#), *Information Retrieval*, **12** (2009), 1–15.
- [12] B. J. Jansen, D. Booth and B. Smith, [Using the taxonomy of cognitive learning to model online searching](#), *Information Processing and Management*, **45** (2009), 643–663.
- [13] C. Kuhlthau, *Seeking Meaning*, 2nd ed., Libraries Unlimited, Westport, CT, 2004.
- [14] S. K. MacGregor and Y. Lou, [Web-based learning: How task scaffolding and web site design support knowledge acquisition](#), *Journal of Research on Technology in Education*, **37** (2004), 161–175.
- [15] G. Marchionini, Exploratory search: From finding to understanding, *Communications of the ACM*, **49** (2006), 41–46.
- [16] G. Marchionini and H. Maurer, [The roles of digital libraries in teaching and learning](#), *Communication of the ACM*, **38** (1995), 67–75.
- [17] J. Ormrod, *Educational Psychology? Developing Learners*, 7th Ed., Pearson, New York, 2011.
- [18] S. Y. Rieh, K. Collins-Thompson, P. Hansen and H. J. Lee, [Towards searching as a learning process: A review of current perspectives and future directions](#), *Journal of Information Science*, **42** (2016), 19–34.
- [19] J.-L. Shih, C.-W. Chuang and G.-J. Hwang, An Inquiry-based Mobile Learning Approach to Enhancing Social Science Learning Effectiveness, *Educational Technology & Society*, **13** (2010), 50–62.
- [20] P. Vakkari, [Searching as learning: A systematization based on literature](#), *Journal of Information Science*, **42** (2016), 7–18.
- [21] C. J. Van Rijsbergen, *Information Retrieval* (2nd ed.), Butterworth, 1979.
- [22] A. Walraven, S. Brand-Gruwel and H. P. A. Boshuizen, How students evaluate information and sources when searching the World Wide Web, *Web for information. Computers & Education*, **52** (2009), 234–246.
- [23] T. Willoughby, S. A. Anderson, E. Woodc, J. Mueller and C. Ross, [Fast searching for information on the Internet to use in a learning context: The impact of domain knowledge](#), *Computers & Education*, **52** (2009), 640–648.
- [24] C. Yin, H.-Y. Sung, G.-J. Hwang, S. Hirokawa, H.-C. Chu, B. Flanagan and Y. Tabata, Learning by Searching: A Learning Environment that Provides Searching and Analysis Facilities for Supporting Trend Analysis Activities, *Educational Technology & Society*, **14** (2013), 1865–1889.
- [25] X. Zhang, J. Liu, C. Liu and M. Cole, Factors influencing users' perceived learning during online searching, in: *Proceedings of the 9th International Conference on e-Learning (ICEL-2014)*, 2014, 200–210.

Appendix A. Search topics. The tasks in the user experiment asked participants to find and save as many relevant documents as possible for answering the topic questions. The topics were presented unchanged from the TREC Genomics Track

descriptions. The topics, with TREC topic numbers noted, as presented to the participants were:

- 2 Generating transgenic mice:** Need: Find protocols for generating transgenic mice. Context: Determine protocols to generate transgenic mice having a single copy of the gene of interest at a specific location.
- 7 DNA repair and oxidative stress:** Need: Find correlation between DNA repair pathways and oxidative stress. Context: Researcher is interested in how oxidative stress affects DNA repair.
- 45 Mental Health Wellness-1:** Need: What genetic loci, such as Mental Health Wellness 1 (MWH1) are implicated in mental health? Context: Want to identify genes involved in mental disorders.
- 42 Genes altered by chromosome translocations:** Need: What genes show altered behavior due to chromosomal rearrangements? Context: Information is required on the disruption of functions from genomic DNA rearrangements.
- 49 Glyphosate tolerance gene sequence:** Need: Find reports and glyphosate tolerance gene sequences in the literature. Context: A DNA sequence isolated in the laboratory is often sequenced only partially, until enough sequence is generated to identify the gene. In these situations, the rest of the sequence is inferred from matching clones in the public domain. When there is difficulty in the laboratory manipulating the DNA segment using sequence-dependent methods, the laboratory isolate must be re-examined.

E-mail address: ae9101@wayne.edu