Evaluating user search trails in exploratory search tasks

1. Exploratory searching behavior is characterized by significant uncertainty towards the goals of a search or a dearth of knowledge about a search topic’s domain (White and Roth, 2009, pp.10; Kuhlthau, 2004).

与事实查询和已知项搜索不同，探索性搜索usually spans across many facets and sub tasks that require multiple queries and search sessions before an information searcher can find meaningful and useful information that extends their knowledge about the task at hand.

2. 评估是否需要帮助：search process prediction model：whether a user is likely to under-perform in an exploratory search task

3. 给予他人检索的推荐：recommend a search trail based on other users search behaviors in a similar context

4. 对推荐效果的评估evaluate：

①use Open Directory Project (ODP)-based categorization of user- traversed Web pages to evaluate each user s information coverage；结构树，深度和广度

②evaluate the order of users search trails while simultaneously incorporating a novel set of metrics that use adjacency of queries issued and Web pages traversed

5. 动态推荐：Evaluating a user s profile in a dynamic manner in order to provide real time recommendations to improve their search performance can be considered an improvement upon traditional recommender systems

当前IR系统的弊端：use query auto-completion and related services to provide recommendations，但是some exploratory searchers may not even have a reasonable query to begin with

6. 目的：attempt to evaluate recommendations based on dynamic user behavior analysis throughout the search process in a specific order of search paths and visitations

7. search trail：a search path that begins with a search engine query and comprises a set of pages visited until the trail terminates with a new query or an inactivity or timeout.( White and Huang, 2010)

Trail selection methods could discount trails with numerous cases of rapid backtracking, or they could maximize rele- vance, coverage, diversity, novelty, and utility by recommending only high quality trails(一是直接推荐高质量trail). Alternatively, we can personalize trail recommendation by weighting trails based on users re-finding behavior（二是以用户为中心推荐个性化推荐）. We can also perform a priori trail analysis to recommend trails when the destination is unclear ( users end up on many pages), and present trail destinations when the destination is clear (many users end up at the same page)（视情况而定）.

8. 用户满意度挖掘：a personalized approach to search behavior analysis is important in order to develop tailored models that provide recommendations to satisfy each individual user.

In order to facilitate search systems ability to understand users, many researchers have taken two viewpoints - one to provide sequential pattern analysis and the other to build behavioral models

9. Egusa et al. (2010) proposed a user-centered method to evaluate the performance of an exploratory search task by comparing the users mental representation of the topic using concept maps. They were able to show empirically that the concept maps between the pre- and post-searches indicated that the users significantly changed their knowledge structure of a topic by completing the exploratory search task.

10. 实验流程

实验情境：实验室、课堂、现场

收集的事务日志数据包括时间戳、用户搜索操作、查询、web页面、书签和用户认为相关的代码片段

11. 实验结果

①introducing new sub topics can act as a way of improving the exploration and knowledge discovery of users；early identification of struggling and recommendation of search trails leads to higher information gains

②发现的可能性：recommended order of search trails surpasses both random and actual order of search trails in terms of LD measure. Again, this emphasizes that the recommended search trail order ensures hard-to-discover pages are found first.

检索式顺序：once a user has gathered information throughout their search processes, sites that have the most clicks would yield a higher goodness measure when their information is deemed useful by an increasingly better informed user.

网页顺序：也很重要

12. 讨论：We showed that the information coverage and the order of the recommended queries and pages at the end of each search trail are more important than merely providing a set of useful or relevant Web pages to the under-performing users.

问题：探索性的搜索任务主要涵盖一般知识，而不需要高度专门化的领域专业知识

Building the Trail Best Traveled: Effects of Domain Knowledge on Web Search Trailblazing

1.①How users blaze trails in a Web search context；②Similarities and differences in how domain experts and novices blaze trails, and if and how the trails eventually generated by the two groups differ

2. Guided Tours继续看：Zellweger [37] introduced scripted documents that were more dynamic than guided tours because they included conditional and programmable paths, auto-mated playback, and active entries.

Chalmers et al. [10] proposed that human recommenders construct and share Web navigation paths.

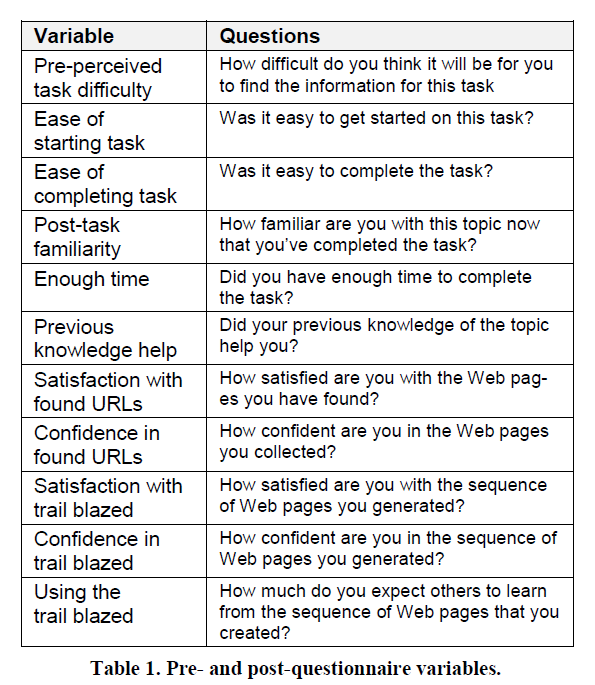
Wexelblat and Maes [28] introduced annotations in Web browsers called footprints that reveal trails through a Web site assembled by the site s designer.

Guinan and Smeaton [13] generated a tour for a given query based on term-matching for node selection and inter-node relationships (e.g., is\_a, precedes) for node ordering.

Wheeldon and Levene [29] proposed an algorithm for generating trails to assist in Web navigation.

3. Step-at-a-Time Recommendations：和传统的推荐有什么不同，继续看ScentTrails [17] combined browsing and searching into one interface by highlighting potentially valuable hyperlinks

4.pre-task: topic familiarity, topic expertise, and perceptions of pre-task difficulty



5. trailblazing process

从检索式看：Previous work has shown that the Web is fertile ground for those without medical training to become unduly concerned about serious medical conditions [33].业余者更多关注癌症而专家更加关注解决办法，This underscores the importance of domain expertise when searching in a sensitive topic area such as healthcare, and provides a good use case for when trails may have a positive impact on search decisions.

从URL看：专家多使用政府和教育网站，是为了帮助非专业人士理解

轨迹：专家筛选掉了更多的URL experts found more sophisti-cated information, but filtered it for inclusion in the trails.

时间：专家更快

专家和普通人的Perceptions of the Trailblazing Process显著不同



专家的URL更具相关性：we obtained human relevance judgments for over one hundred thousand queries that were randomly sampled by frequency from the query logs of the Bing commercial search engine.

6. trail outcomes

网站：domain-specific medical sites were more popular than search engines among experts. The most popular domain for novices was google.com

信息来源和展现方式会影响网站的选择：Experts selected pages providing summary and structured information while novices favored content-rich pages

搜索轨迹结构：





搜索轨迹价值：

被试自评，标准化

6. The information gathered during this study can also be used to inform the development of automatic trail generation algorithms e.g., [19,29]. Following experts’ example, the algorithms could require that trails begin with background information and become more specific as the trail progress-es, be arranged in chains rather than as hubs and spokes, and favor structured overviews over detailed content.

7. 专家与新手虽有差异但是差异不大；领域特异性

Assessing the Scenic Route: Measuring the Value of Search Trails in Web Logs

1. Log data：These log entries include a unique identifier for the user, a timestamp for each page view, a unique browser window identifier (to resolve ambiguities in determining which browser a page was viewed), and the URL of the Web page visited（两个月）

2. trail mining：

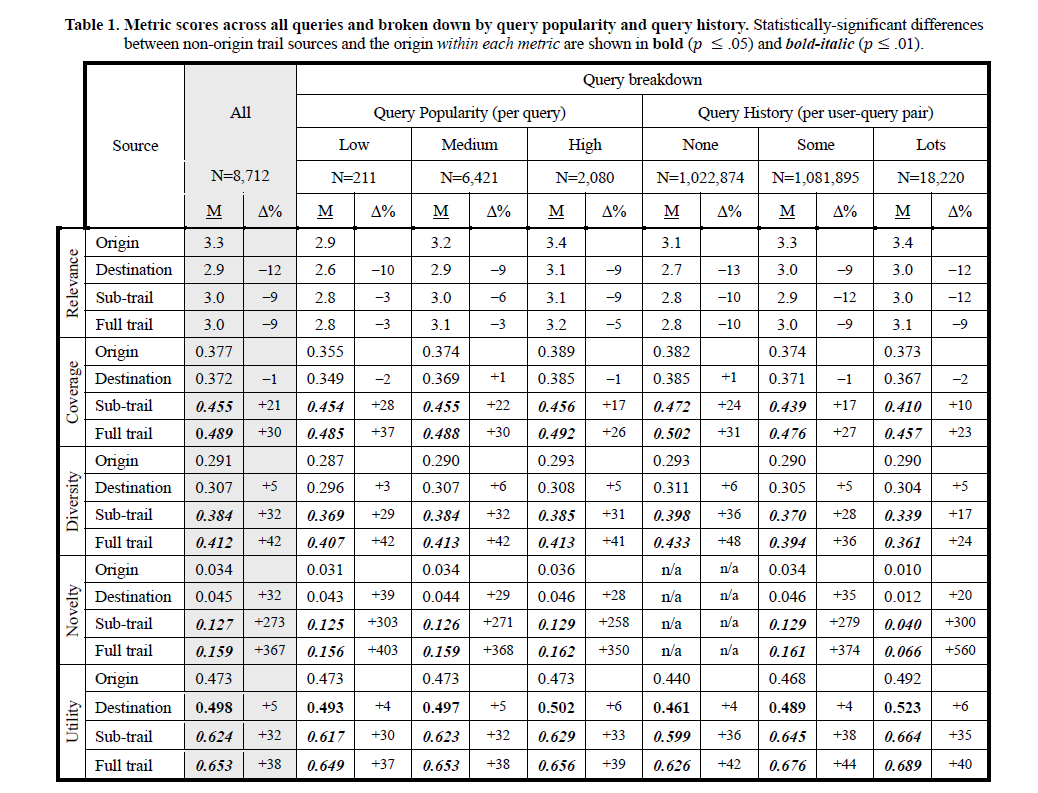


3. Trail Labeling：coverage, diversity, and novelty

we classified the Web pages sourced from each context into the topical hierarchy from a popular Web directory, the Open Directory Project (ODP) (dmoz.org)

4. 研究问题：four sources (origin, destination, sub-trail, and full-trail), which: (i) provide more relevant information? (ii) provide more topic coverage? (iii) provide more topic diversity? (iv) provide more novel information? (v) provide more useful information?

5. 指标relevance, coverage（占标准范围的比例）, diversity（label数量）, novelty（与以前不同的label的数量） and utility



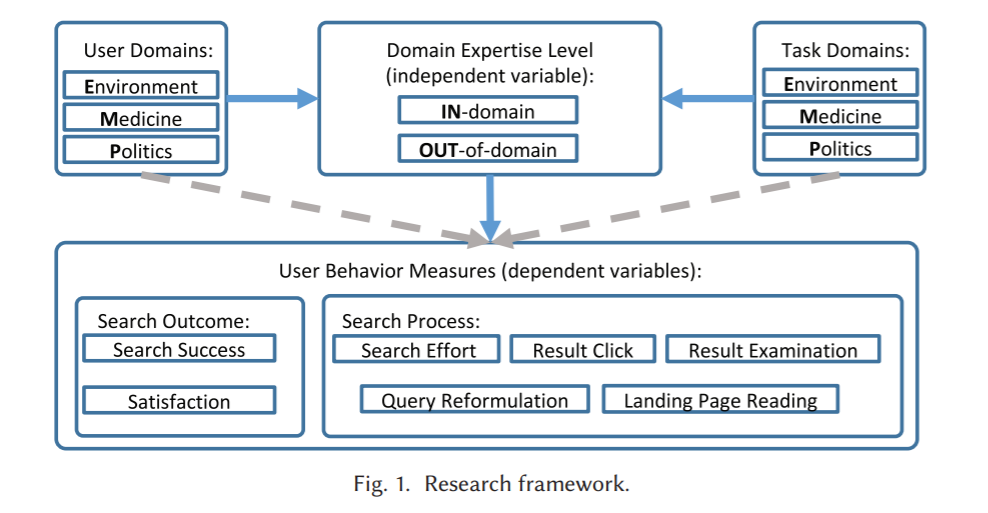
Mao, J., Liu, Y., Kando, N., Zhang, M., & Ma, S.(2018). How does domain expertise affect users' search interaction and outcome in exploratory search?. *Acm Transactions on Information Systems,* *36*(4), 42.1-42.30.

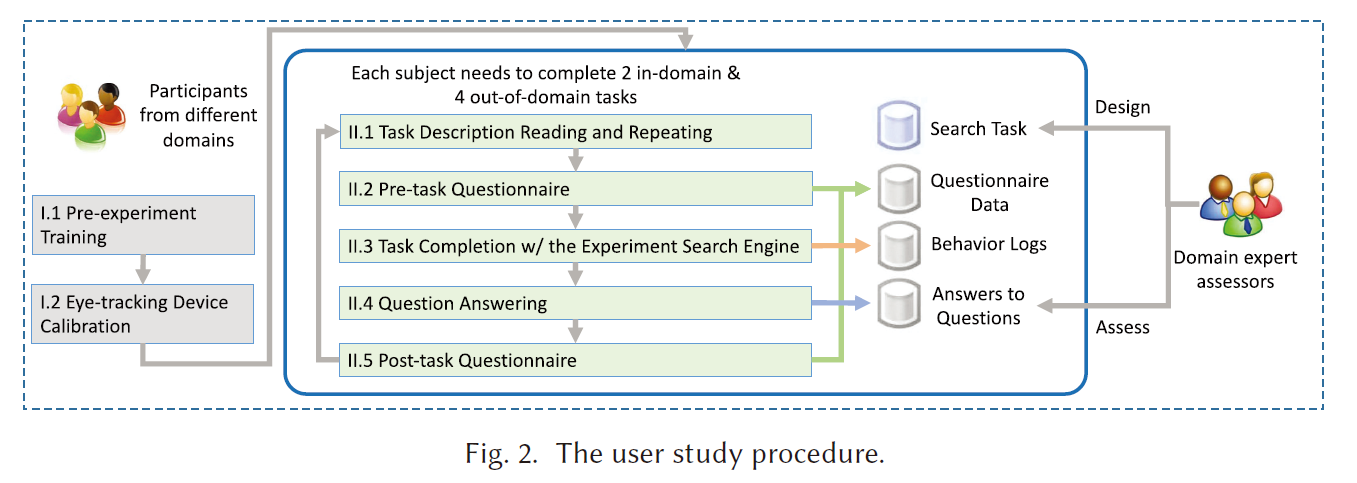
1. RQ1: What is the relationship between domain expertise and search outcomes?

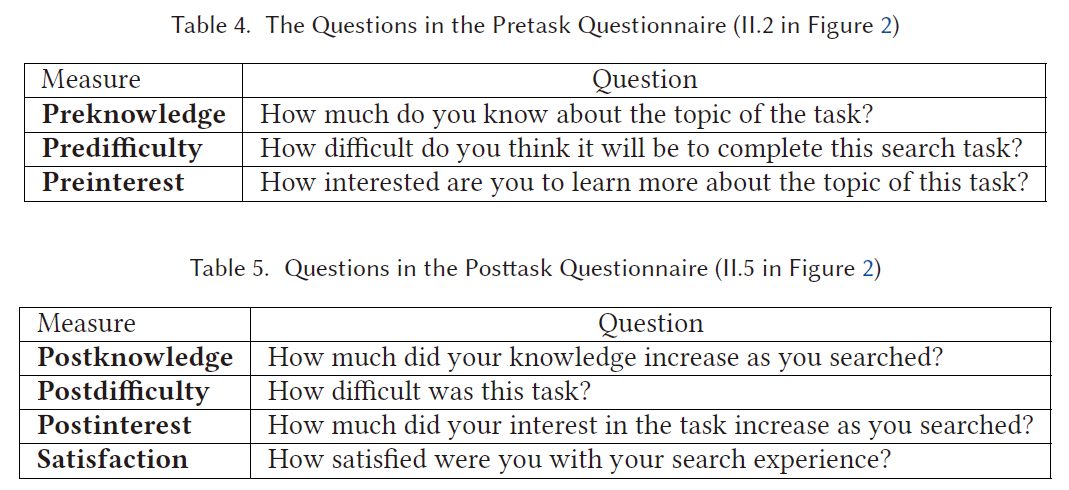
RQ2: Which user behaviors, including how the user formulates queries, interacts with SERPs, and reads the landing pages, will be affected by domain expertise in exploratory search?

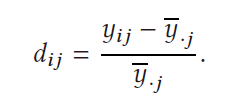
RQ3: Are the influences of domain expertise consistent across different knowledge domains?

2.







3. 亮点：control the task and domain effects 

4. declarative domain knowledge（用户关于搜索的主题领域的知识）and procedural domain-specific search strategies(特定领域的程序性的搜索策略)

5. ①Our analysis confirms that a high domain expertise level often leads to a higher success rate in completing the search tasks, but we fail to detect a domain-independent domain expertise effect on users satisfaction；

②with task descriptions being the major source for query vocabulary, the participants may use more new query terms from landing pages and SERPs when exploring unfamiliar domains;

③ We find that the results clicked by domain experts may not be more relevant (Section 4.3.3). Therefore, we should rethink about how to exploit domain experts click logs.

④With an eye tracker, we find that the nonexpert users use more terms encountered during search as their query terms, especially in the medicine domain. We should enhance the query suggestion function for highly technical domains like the medicine domain because exploring such domains may require domain-specific query vocabulary.

(3) Regarding RQ3, because the effects of domain expertise level on users search behavior may be different in different knowledge domains, when trying to personalize the search results according to the user s domain knowledge level, we need to develop different models for different knowledge domains

Liu, J., & Zhang, X.(2019). The role of domain knowledge in document selection from search results. *Journal of the Association for Information Science and Technology*, *70*(11), 1236-1247.

1.

RQ1. Do people with different levels of DK select the same or different set of documents for the same search topic?

RQ2. For the selected documents on the search result pages (SERPs), are there differences between DK experts and novices regarding the documents ranking positions

RQ3. Are there differences between domain experts and novices in assessing the relevance of their selected documents, including the documents that were selected by both groups

RQ4. Do the documents selected by domain experts have different features than those selected by domain novices

2.

The study designed an experimental search system using the Indri search engine from the Lemur toolkit.2 The system s underlying data set used a subset of the MEDLINE bibliographic database (Hersh &amp; Voorhees, 2009) in the Text Retrieval Conference s (TREC3) 2004 Genomics track data collection. This subset was for the period of year 2000 2004 (n = 1.85 million), the size of which was large enough to allow reasonable retrieval efficiency

The topics were selected based on the consideration of balancing (i) MeSH categories, (ii) the specificity levels of the topics, and (iii) search difficulty levels.

(ii)Task specificity was determined by the level of the task topic keywords in the MeSH tree, specifically, the path length to the root in the MeSH category tree, as judged by an external expert in the biomedical area hired in the study. A topic having a MeSH hierarchy level of higher than three was considered as a general, and otherwise, specific.

A participant s DK level was determined by two factors: (i) his/her familiarity rating for the MeSH terms, and (ii) his/her search topic familiarity and expertise ratings in the pretask questionnaires.

3.

Although some common documents (less than a quarter of all documents) were selected by both high and low domain knowledge users, the two groups selected different sets of documents, and the number of documents read by domain experts was higher than for novice users

domain novices tended to view documents ranked on the top of the SERPs (that is, those deemed more relevant by the systems)

domain experts were found to have higher self-assessed relevance scores, and their selected documents also had higher relevance scores assessed by the TREC gold-standard evaluation

domain experts selected documents that were longer, had more MeSH terms, and more general MeSH terms.

O'Brien, H. L. , Kampen, A. , Cole, A. W. , & Brennan, K. . (2020). The Role of Domain Knowledge in Search as Learning. CHIIR '20: Conference on Human Information Interaction and Retrieval.

1. 不同点：Search as Learning 搜索策略

2. Participants who are less familiar with search topics may achieve greater knowledge gains. The search may have turned up more new information for NE than for DE.

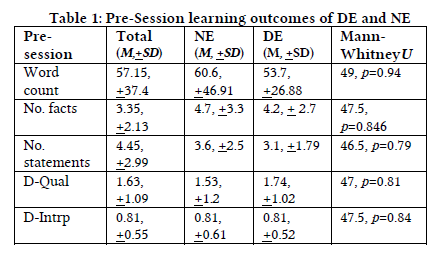
With the exception of the fact-finding task, DE self-reported greater learning gains than NE。

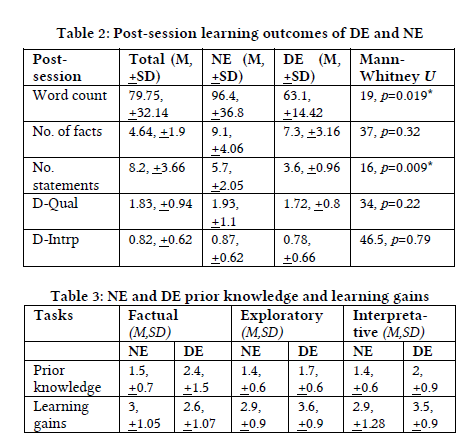
General searchers wrote longer summaries than domain experts, but there were no significant differences in the learning outcomes between the two groups.

3. White et al. [23] found differences in DE s and NE s search behaviours. Specifically, DEs issued longer queries with more domain-specific vocabulary terms; spent more time searching and visited more pages; and drew upon more technical sources.

Ryen W. White, Susan T. Dumais, and Jaime Teevan. 2009. Characterizing the influence of domain expertise on web search behavior. In Proceedings of the Second ACM International Conference on Web Search and Data Mining – WSDM’09, ACM Press, Barcelona, Spain, 132. DOI: https://doi.org/10.1145/1498759.1498819

4. gains是被试的主观判断





Hsieh-Yee, I. . (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science, 44*(3), 161-174.

Data analyses showed that subject knowledge interacted with search experience, and both variables affected searchers behavior in four ways:

(1) when questions in their subject areas were searched, experience affected searchers use of synonymous terms, monitoring of the search process, and combinations of search terms;

(2) when questions outside their subject areas were searched, experience affected searchers reliance on their own terminology, use of the thesaurus, offline term selection, use of synonymous terms, and combinations of search terms;

(3) within the same experience group, subject knowledge had no effect on novice searchers; but——

(4) subject knowledge affected experienced searcher s reliance on their own language, use of the thesaurus, offline term selection, use of synonymous terms, monitoring of the search, and combinations of search terms.

The results showed that search experience affected searchers use of many search tactics, and suggested that subject knowledge became a factor only after searchers have had a certain amount of search experience.

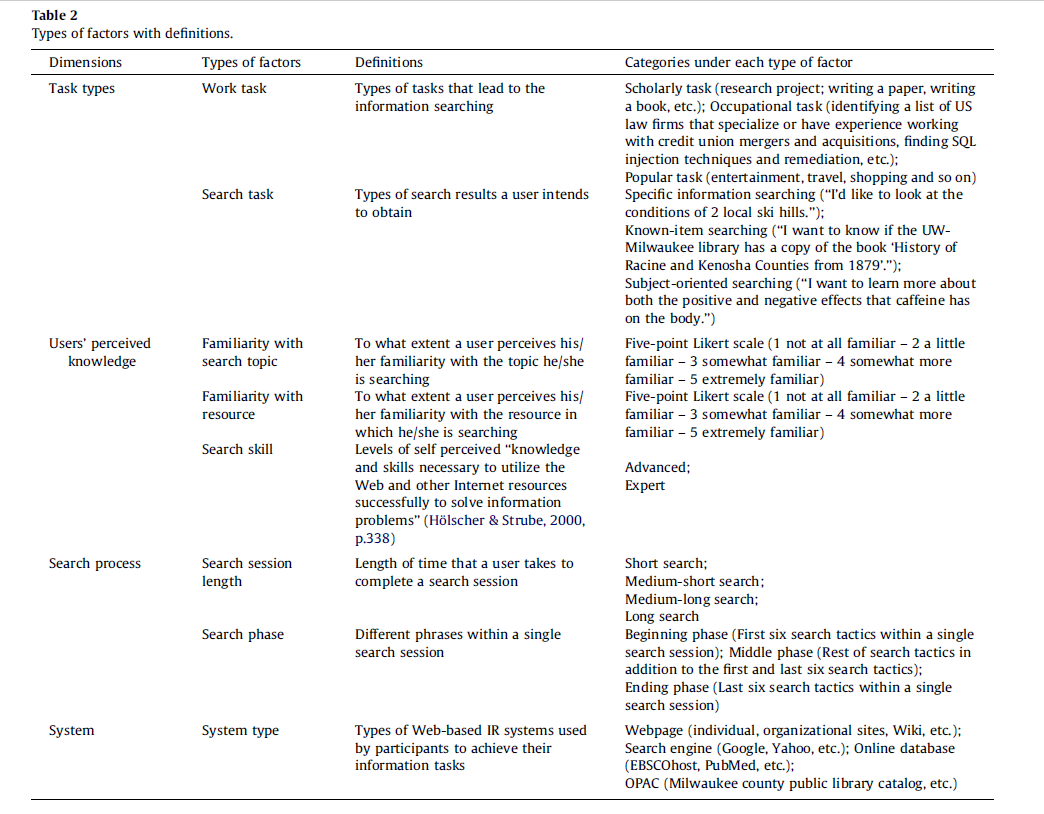
Factors affecting the selection of search tactics: Tasks, knowledge, process, and systems

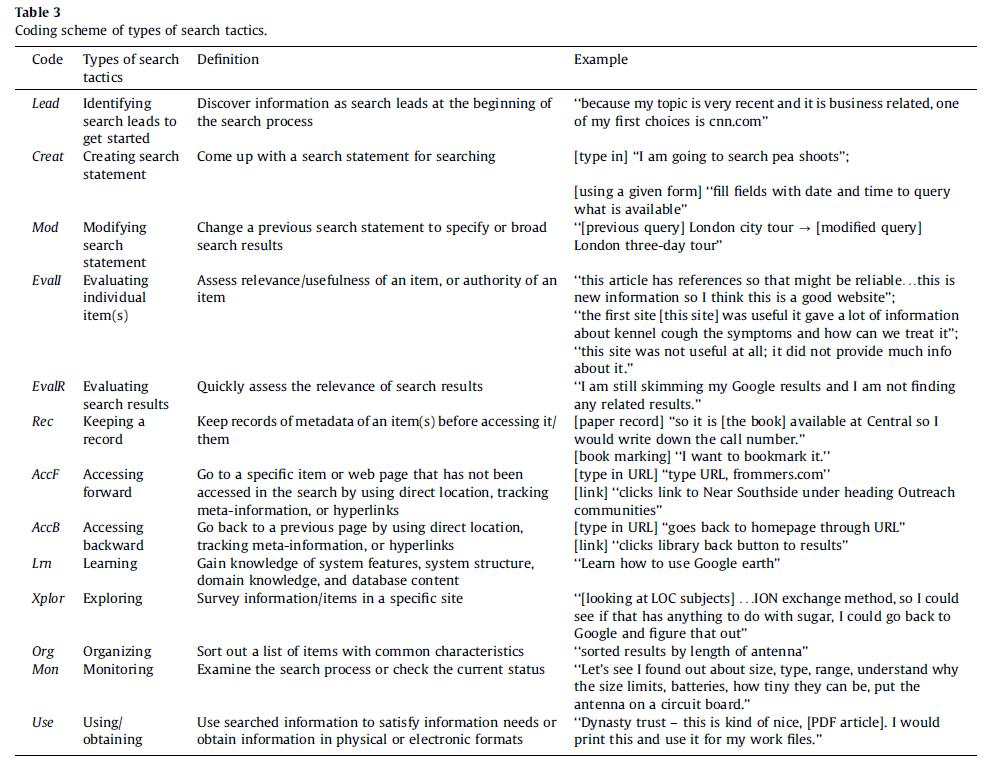
This study investigated whether and how different factors in relation to task, user-perceived knowledge, search process, and system affect users search tactic selection

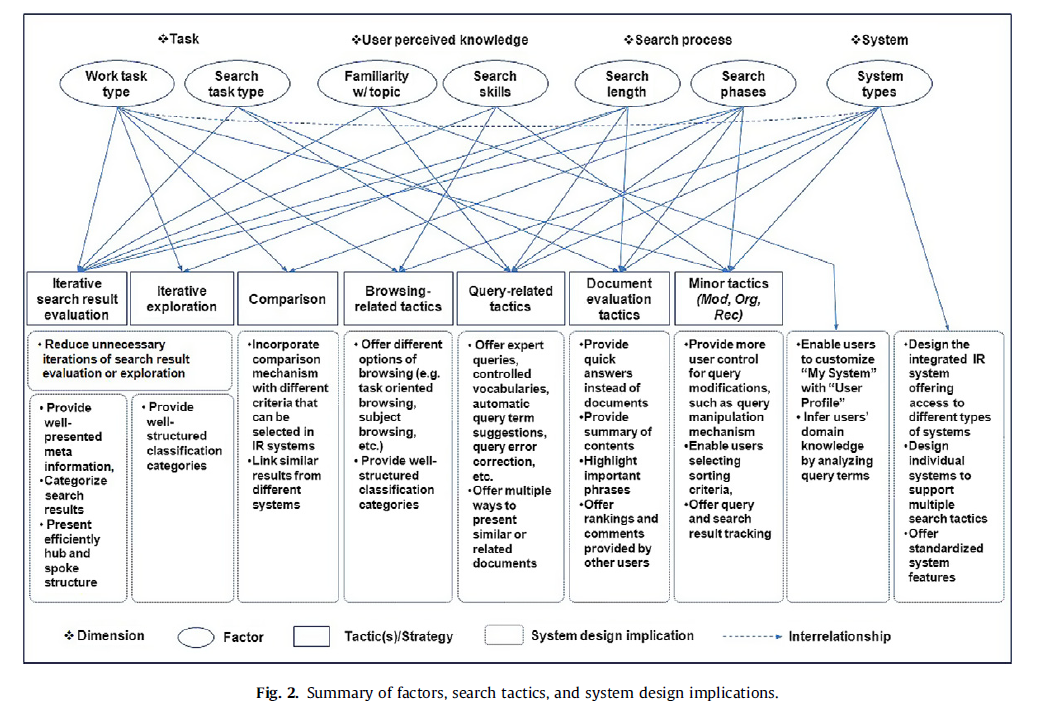
1. procedure

First, participants were instructed to fill in a pre-questionnaire requesting their demographic information and their experience in searching for information. From the **pre-questionnaire**, **data in relation to search task type, work task type, familiarity with topic, and familiarity with source were collected**. Second, participants were asked to **keep an information interaction diary for 2 weeks to record the process of achieving one work-related and another personal-related task**. The diaries require participants to **log their source selections, search tactics, and reasons associated with why they applied or did not apply different types of search tactics**. From the diaries, the authors attempted to identify factors behind their selection of search tactics qualitatively. Third, **think aloud protocols and logs were employed to record their search processes and associated verbal protocols**. Participants were invited to **come to the lab to search for information for one work-related and one personal search tasks**. They were instructed to think aloud during their information-seeking process. Their search processes were captured by Morae, a usability testing software that records users movements, as well as capturing their think aloud protocols. Finally, participants were also asked to fill in **post-questionnaires**. Their answers regarding **their selection of information sources, their search tactics, their problems, and factors affecting their selection of search tactics were noted.**

**2.**

****

****

****

**专家/新手：搜索特征的区分**

**Ryen W. White的时间线**

**互联网无法区分用户是否是领域专家，可以根据阅读的特点进行区分**

**搜索之后的感想：众包平台上强迫认真**

**看search trail 的设计**

**用户搜索记录对现在用户搜索的影响**