Information Processing and Management 56 (2019) 1010–1025

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Exploring the immediate and short-term effects of peer advice and cognitive authority on Web search behavior

Jiqun Liu*, Yiwei Wang, Soumik Mandal, Chirag Shah

Rutgers, the State University of New Jersey, 4 Huntington Street, New Brunswick, NJ 08901, United States

A B S T R A C T

An individual's Web search behavior can be influenced by a number of factors, including features and functions of a search engine as well as search education. In contrast to the long-lasting attention to the algorithm and interface dimensions of search, there is a lack of research concerned with the potential effects of user education on search behavior. To address this gap, we ran a three-session field-lab-combined study to examine the effects of user education from two distinct sources – peer advice and cognitive authority (operationalized as video-based student's advice and expert's advice respectively) – on Web search behavior in two different search task scenarios (i.e., factual specific and factual amorphous tasks). We also tested if these behavioral effects persist for a short period of time when the explicit search tips are removed. Using 185 task session data generated by 31 participants in two field and one lab sessions, this study demonstrates that: (1) both peer advice and cognitive authority are effective in stimulating immediate behavioral changes in Web search; (2) the immediate behavioral impact of search advice is broader in factual amorphous task than in factual specific task; (3) framing search tips as the advice from cognitive authority is more likely to generate continuing, short-term effects on Web search behaviors. This research has implications for the design of task-aware user education as well as the study of users' interactions with IR systems in general.

1. Introduction

Users' interaction with Web search engines is one of the major components of ubiquitous online information activities. Research on the interventions on search behaviors often focuses on the algorithm and interface aspects, such as learning-oriented retrieval algorithm (Syed & Collins-Thompson, 2017), size of search engine result page (SERP) (Kelly & Azzopardi, 2015), and innovative search assistant tools embedded in search interfaces (e.g., Capra, Arguello, Crescenzi, & Vardell, 2015). However, changes may also directly come from the user side. Effective user search education can improve searchers' expertise, affect their search behaviors, and thereby shape their interactions with information retrieval (IR) systems (Moraveji, Russell, Bien, & Mease, 2011). Despite the potential value for user-oriented search education, the effects of different types of user education on Web search have been scarcely examined in interactive IR community.

People often interact with other individuals or groups when seeking and searching for information to resolve the encountered gaps in sense making, tasks, or information supply-demand disequilibrium (Liu, 2017; Schutz & Luckmann, 1973). For instance, in everyday life information practices, individuals often ask for peer advice from friends, family members, and colleagues in information interpretation and decision making (Gardner & Steinberg, 2005). One of the contextual factors behind the active solicitation of peer advice is the similarity in social comparison. According to the social comparison theory and related empirical evidences, when making choices, people prefer to compare themselves with others who are similar to them in opinions, background, and abilities (Cohn, Fehr, Herrmann, & Schneider, 2014; Festinger, 1954). Besides, researchers and educators have also identified cognitive authority as a critical factor that affects people's behaviors and judgments to varying extents (Wilson, 1983). In the context of

* Corresponding author.
E-mail address: jl2033@scarletmail.rutgers.edu (J. Liu).

https://doi.org/10.1016/j.ipm.2019.02.011
Received 24 March 2018; Received in revised form 4 February 2019; Accepted 15 February 2019
Available online 22 February 2019
0306-4573/

information seeking and search, suggestions and opinions from cognitive authority can significantly affect users' judgments of information credibility and their usage of information sources in problematic situations (Rieh, 2002). Given the potential impacts from peers and authorities on information behavior, a research question naturally arises in the interactive IR (IIR) context: To what extent do peer advice and cognitive authority affect Web search behavior?

In addition to the design of user education methods and techniques, the temporal aspect of user education effects is also critical for its application in real-life information search and should be properly examined in IR research. Due to people's natural tendency of avoiding cognitive dissonance and the priming effect of previous stimuli, when under the influences of similar contexts and cognitive processes, users may still tend to behave in a relatively consistent and continuous manner (Tanford & Montgomery, 2015; Tipper, 1985). These theories and evidences suggest that the behavioral effect of appropriate user education may be retained at least in a short period of time even when explicit instructions are removed. Thus, after examining the influences of different types of user education and instructional interventions (i.e., peer-advice-based and cognitive-authority-based), a follow-up question would be: If the immediate effects of user education on users' Web search behavior do exist, would they last (at least for a short period of time) when the search context (e.g., physical environment, task) changes?

Given the unanswered questions above, the work described in this paper aims to go beyond the mainstream interface- and task-oriented IR approaches, and investigate the potential influences of different types of user education on Web search. Specifically, the authors sought to *explore the effects of peer advice and cognitive authority on Web search behavior, and to examine if these effects persist when search context varies*. Regarding operationalization, cognitive authority was represented by a search expert (university professor)'s advice on how to search effectively, whereas peer advice was operationalized as two students recommending search tips as peer advisors. Each type of advice was presented to study participants, who were students at a university, using a video.

This study has multiple contributions and implications:

- On a theoretical level, it goes beyond the interface and search task dimensions and extends the current understanding of the influences of peer advice and cognitive authority on the interactions between users and Web search engines in online information seeking episodes.
- Within the scope of search engine user education research, our results demonstrate that among different components of user education, in addition to the content features of search advice (c.f. Moraveji et al., 2011; Yamamoto & Yamamoto, 2018), the difference in the source of user education (i.e., from peer or cognitive authority) can also lead to divergent behavioral effects on some aspects of Web search. In this sense, this study partially reveals the complexity and multidimensionality of search education.
- From system features and instructional intervention design perspectives, this work has implications for IR system designers and search educators on how to leverage the strengths of search advice of different types in guiding users' Web search activities and influencing their search tactics.

2. Related works

2.1. User education in Web search

Although the topic of *user education* has not motivated a wide discussion and exploration at conceptual level within IR community, the idea of user education has already been incorporated and operationalized in a variety of IR research on search assistance and recommender systems design. In IIR user studies, user education is usually operationalized as optimal search tips and advice on query formulation, page browsing, and search path selection (e.g., Moraveji et al., 2011). In previous user-education-related studies, researchers mainly focused on evaluating the content and effectiveness of searching assistances and system recommendations based on search outcomes, such as query term suggestion, search tactic suggestion, and search strategy suggestion (Jansen, 2006; Jansen & McNeese, 2005; Moraveji et al., 2011; Yamamoto & Yamamoto, 2018). Specifically, for example, Moraveji et al. (2011) constructs search tasks that are designed to benefit from the introduced search tips (e.g., use quotation marks around query, use minus sign) in their user study and found that participants who were shown search tips still demonstrate improved search ability and efficiency when the explicit search tips were removed. Yamamoto and Yamamoto (2018) demonstrates that using query terms that stimulate critical thinking (e.g., analysis, evidence, demonstration) can effectively encourage careful query formulation, information seeking, and decision making.

Among various types of user education and instructional intervention design, suggestions for query expansion are the most common class of recommendations and have been widely employed in different search systems, task contexts and scenarios. To generate relevant and useful recommendations for query expansion, researchers have exploited features of different types extracted from documents, relevance feedbacks, and the associated search interaction data (e.g., Cao et al., 2008; Cucerzan & White, 2007; Xu & Croft, 1996). In addition to the system-side and individual-user features, a few researchers also explored the social dimension of query recommendation. For instance, drawing upon classical theories of social influence, Kelly, Cushing, Dostert, Niu, and Gyllstrom (2010) examined the effects of popularity and quality on the actual usage of query recommendation and found that quality was more important than popularity. This line of research demonstrates that factors from social dimensions should be taken into account in user education design and offers an alternative, social-context-based approach to exploring the effect and design of IR user education from the user side.

In contrast to query suggestion, suggestions of search tactics and strategies refer to sequences of search actions or instructional search statements, rather than a single component or moment of search (Belkin, Marchetti, & Cool, 1993; Brajnik, Mizzaro, Tasso, & Venuti, 2002). The major difference between them is that search tactics are relatively shorter sequences specified for solving a single

information goal or intention, whereas search strategies are longer sequences of search actions generated and recommended for solving multiple goals or complex search tasks (Bates, 1979; Xie & Joo, 2010). In Web search, tactical recommendations can provide immediate guidance to exploring unfamiliar information sources, patches, and visualizations without significantly increasing the cognitive load on query reformulation (Moraveji et al., 2011). Search strategy suggestions can facilitate complex task completion or even proactively help users improve their search strategy in real time and avoid potential problems (Kriewel & Fuhr, 2007; Shah, 2018). While previous research on tactical and strategy suggestions enhances our understanding of the underlying relationships between task, system features, and recommendations design, the role of the social dimension of search tactics and strategy suggestions still remains ambiguous. This lack of exploration on the social dimension may create a bottleneck for IR user education studies as users' motivations and interactions with information and systems often involves factors from their daily social interactions.

2.2. Social comparison and cognitive authority

When seeking for information to resolve problematic situations, people usually seek advice and feedbacks from both peers and cognitive authority. The social effect of peer advice has been confirmed and articulated in social comparison theory. According to Festinger (1954)'s central propositions on social comparison, people prefer to compare themselves with others who are similar to them in background, opinions, culture, and abilities. For example, college students may be more likely to compare their Web search skills and strategies with other college students, rather than with experts and professors in the area of Web search. Hence, peer advice may lead to behavioral effects in Web search education as people prefer to take cues from those whose experiences resemble theirs. While the role of similarity in social comparison was demonstrated in a variety of field experiments (Cohn et al., 2014; Frey & Meier, 2004; Suls, Martin, & Wheeler, 2002), it has not been fully understood in the context of user-IR system interactions.

Despite the power of peer advice, people are also likely to solicit knowledge from cognitive authority in making judgments, especially in the context of medical information acquisition and evaluation (Wathen & Burkell, 2002). Wilson (1983) first defined the term "cognitive authority" to differentiate the authority in information interpretation and knowledge construction from "administrative authority," which refers to the widely-acknowledged right of a person who holds a certain position. In the context of online information seeking, since people usually look for topically relevant and authoritative information, the individuals, groups, and other nonhuman sources (e.g., official websites) recognized as cognitive authorities are often regarded as reliable sources of second-hand knowledge (Rieh, 2002; Wilson, 1983). Although cognitive authority has potential effects on Web search, its implications for IR recommendation system and search education design have been scarcely discussed. In addition, while user education research has discussed the inclusion of information literacy skills in curriculum development (Kafai & Bates, 1997; Scott & O'Sullivan, 2005), the video-based educational content as immediate intervention has rarely been studied. To determine a suitable social approach for instructional intervention design and better understand the social dimension of user education, the behavioral effects of peer advice and cognitive authority needs to be compared and explored in depth.

2.3. Methods of Web search studies

Methodologically, a majority of the existing research on Web search behavior can be classified into three categories: controlled lab studies, field studies, and large-scale log-based studies (Grimes, Tang, & Russell, 2007). Within each line of study, researchers have defined or employed a variety of behavioral measurements (e.g., querying, page-viewing behavior), aiming to detailedly characterize users' interactions with IR systems in varying scenarios.

Since lab study design is relatively effective in controlling irrelevant contextual variables (Grimes et al., 2007), it has been widely employed to infer the causality between Web-search behavior and task-related factors, such as task complexity, task difficulty, and perceived time constraints (e.g. Jiang, He, & Allan, 2014; Kelly, Arguello, Edwards, & Wu, 2015). Despite the advantages and contributions, lab study design alone is not enough to support a comprehensive exploration of the impacts of the social dimensions of user education as participants are not in a realistic environment using the devices that are actually available to them (Kelly, 2009). Education and intervention designed in lab-based user study often happens in an idealized, controlled environment and thus lacks realism (Grimes et al., 2007).

In addition to the controlled lab study, field study has served as an alternative approach for capturing Web-search behavioral patterns as it allows participants to interact with search engines and tasks in naturalistic settings (Greifeneder, 2016; Kelly & Belkin, 2004; Teevan et al., 2004). In a variety of recently conducted field studies, researchers employed browser plugins of different types in data collection and investigated the interactions between search behavior, search task, and user features (e.g., He & Yilmaz, 2017; Wang et al., 2017). Browser plugins automatically collect participants' natural search activities (e.g., queries, pages) and thus enable researchers to conduct relatively large-scale, long-distance (e.g., crowdsourcing) user studies in naturalistic settings. Despite the advantages, field-study design also has limitations as researchers have little control of various implicit factors that may affect participants' behaviors. In particular, for browser-plugin-based field studies, search behavior data alone may not be able to fully reveal the complexity and variations in users' task contexts and search environments. Consequently, it is often difficult for researchers to differentiate the main effects of the variables of interest (e.g., task type, user characteristics) from the mixed effects caused by irrelevant contextual factors. To compensate for this limitation, researchers often employ a series of qualitative methods (e.g., semi-structured in-depth interview, search diary), aiming to better explore the underlying problematic situations that contextualize users' search interactions.

Since lab study and field study have their respective strengths, combining them within an integrated research design may enable researchers to not only infer the causality between search behavior variations and different types of user education and instructional

design, but to explain the potential lasting effect of user education in naturalistic settings. While large-scale search log studies can also measure users' natural search behaviors and collect rich behavioral data from the backend of search engines (Lucchese, Orlando, Perego, Silvestri, & Tolomei, 2013), they usually lack user characteristics and contextual information and thus may not fit well in studying the behavioral impact of peer advice and cognitive authority.

2.4. Summary

User education has been operationalized as recommendations of query, search tactic, and search strategies in a variety of systems and search contexts. Previous research mainly focuses on the connection between the contents of recommendations and the features extracted from documents and interaction data. While a few researchers explored the social dimensions of user education and IR recommendation (e.g., Kelly et al., 2010), the effects of peer advice and cognitive authority have not been systematically examined in the context of Web search education. By exploring these potential effects in a hybrid, lab-field-combined user study, researchers can embrace a better understanding of users' interactions with information in various scenarios and thereby can leverage the knowledge in future user education and search instructional design.

3. Research questions

Given the research gaps discussed above, we explored the social dimension of search engine user education. Specifically, this study focused on two potential factors from the social dimension of user education, *peer advice* and *cognitive authority*, and sought to understand their behavioral effects in user education. This study aimed to answer three research questions (RQs):

- **RQ1**: What are the effects of peer advice on users' Web search behavior?
- **RQ2**: What are the effects of cognitive authority on users' Web search behavior?
- **RQ3**: Do the effects of peer advice and cognitive authority on Web search behavior last in a short period of time?

To answer these three questions, we designed and conducted a lab-field-combined three-session study that is described in the following method section.

4. Method

4.1. Participant

This study recruited 36 undergraduate students from a variety of disciplines via Rutgers School of Communication and Information email lists and Rutgers Facebook groups. Thirty-one students completed the entire study and their search data were used in our data analysis. In order to be eligible, participants had to be (a) at least 18 years old, (b) fluent in English speaking and writing, and (c) regular Chrome browser users (This study used a Chrome browser plugin to collect log data in all three sessions. Switching browsers may affect participants' regular search behavior). The sample was composed of students from various grade levels and areas of study. We filtered out the search data collected from five of the participants (two from peer advice group, two from cognitive authority group, one from control group) due to system errors and incomplete data collection (some of them did not finish the whole study) and our final sample consists of 31 participants. Seven out of the 31 participants were from Information Technology and Informatics programs, and none of the participants was a Library and Information Science (LIS) student. Most of the participants were native English speakers (77%) and female (68%). They ranged in age from 18 to 24 with an average age of 20.5 (four participants are 18 years old, six participants are 19, seven participants are 20, five participants are 21, six participants are 22, one participant is 23, and two participants are 24). Each participant was compensated $40 in cash after finishing the entire study. To encourage participants to take our study seriously, they were told at the beginning that the top five performers judged by us would receive a $20 bonus. Note that we judged participants' search performances qualitatively based on the answers submitted by them and that no quantitative measure was developed to formally evaluate search outcomes. This is because in this study we only focused on the behavioral impacts of user education. The search-related tips and persuasions (will be introduced later) were not designed to improve search performances. The qualitative performance judgment conducted here was *only* for deciding the extra bonus. For the performance evaluation, we only looked at how detailed and organized their answers were, instead of focusing on if the answers were correct or not. All participants were informed of this criterion after they registered for the study. In other words, we gave out the bonus based on the extent to which they took this study seriously.

4.2. Study procedure

This three-session study was composed of two field sessions and one lab session. It started and ended with a field session in which participants finished two search tasks on devices of their choice at their own locations. Between the two field sessions was the lab session in which participants were randomly assigned to one of three different conditions and performed two search tasks in a computer lab. A combination of log data, diary data, and interview data was collected to depict participants' search experiences from multiple angles and to allow data triangulation. This study was approved by Rutgers Arts and Sciences Institutional Review Board (Protocol #: E 17-399). The detailed procedure is described as follows.

4.2.1. Field session 1

After each participant signed up for the study through a link provided in the recruitment message, a confirmation email was sent to them specifying the study procedure. They followed the instructions to install a Chrome plugin and used a unique username and password combination to log in to the study system. Two types of events were recorded by the Chrome browser plugin, each associated with a unique timestamp, user ID (randomly assigned at the registration stage), task ID, session ID, a URL, and a tab ID. The events included: (1) Search events: query issued by the user and the search engine(s) used. (2) Browsing events: page type (SERP or content page) and actions (i.e., click, bookmark). With respect to participants' control over the data and privacy issues, participants could log out from the plugin when they did not want their search data to be collected.

Participants first filled out a demographic questionnaire asking about their age, gender, educational background, and English proficiency. They then started to search and to detail their search experiences and answers to the task questions in an online diary. Specifically, participants were required to evaluate their perceived level of task difficulty using a 7-point Likert Scale and briefly describe their search strategies, reasons behind their strategies, and what they found out (i.e., their responses to the tasks). The diary is arranged to capture participants' thoughts and experience during the search process and to allow us to qualitatively investigate their search behavior. Participants did not need to be specific about what sources or queries they used in their diary as those were captured by a browser plugin. In addition to the required questions, a text box was provided to allow participants to add anything that was not asked in the diary. The second task followed with the same procedure.

Field sessions aimed to capture individuals' natural search behavior by reducing artificial constraints and manipulations. Therefore, although the two tasks would only take up to an hour in a controlled lab setting, each participant was given up to three days to finish the tasks. They could leave the tasks and log out in the middle, and they would be brought back to their previous stopping point when logging in again. Compared to the lab session, a field session offered a flexible set-up which can better approximate real-life search environments.

4.2.2. Lab session

Each participant came to a computer lab individually to finish two search tasks on a desktop computer provided by us. Each person was randomly assigned by the study system to one of the three groups receiving different treatments (i.e., search education videos): control group (no video/search education), cognitive authority group (search tips operationalized as expert advice), and peer advice group (search tips operationalized as student advice).

Before the actual search, participants in the cognitive authority group watched a 10-min video in which a professor from the field of Information Science introduced tips on how to search on Google while performing two search tasks on a screen. He started with introducing himself in the video (as a professor in Information Science) and both his face and the computer screen were shown in the video. The professor performed two search tasks and explained his search intentions and the search tips he used via thinking aloud. Participants could see the whole search process as well as the associated search tips in the main screen and also see the professor's face in the small screen at the bottom right of the video screen (see Fig. 1). Since this study was conducted in a university

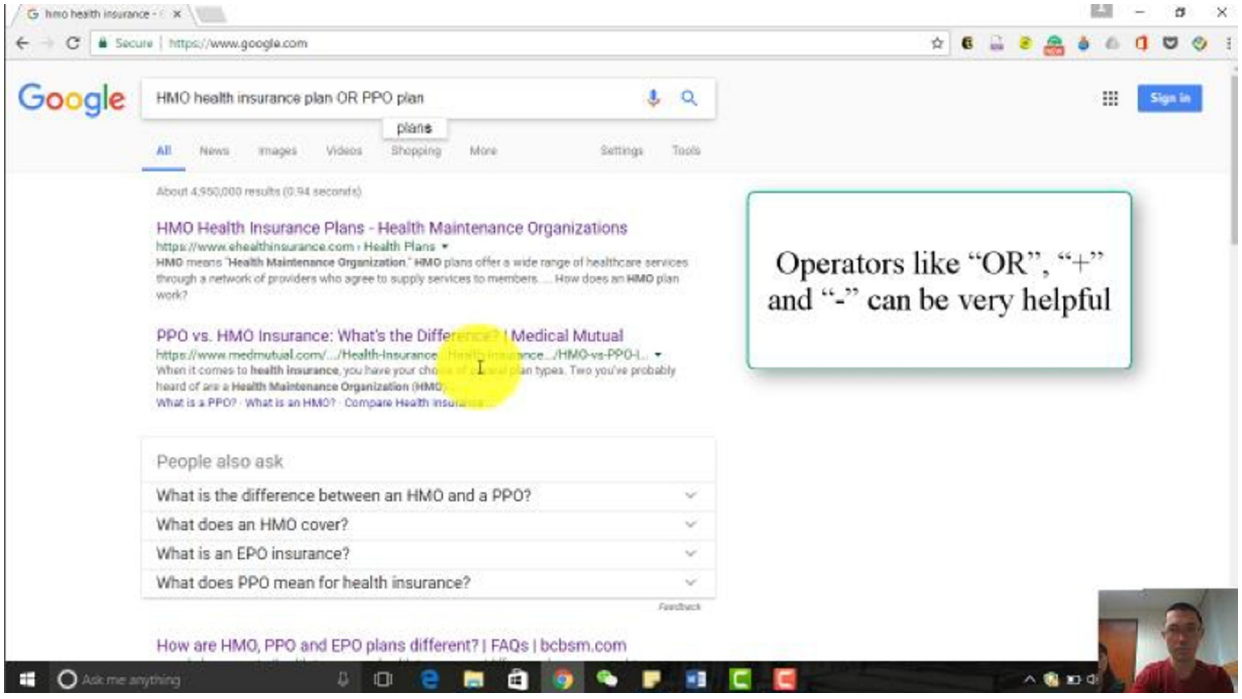

Fig. 1. Video screenshot of peer advice group.

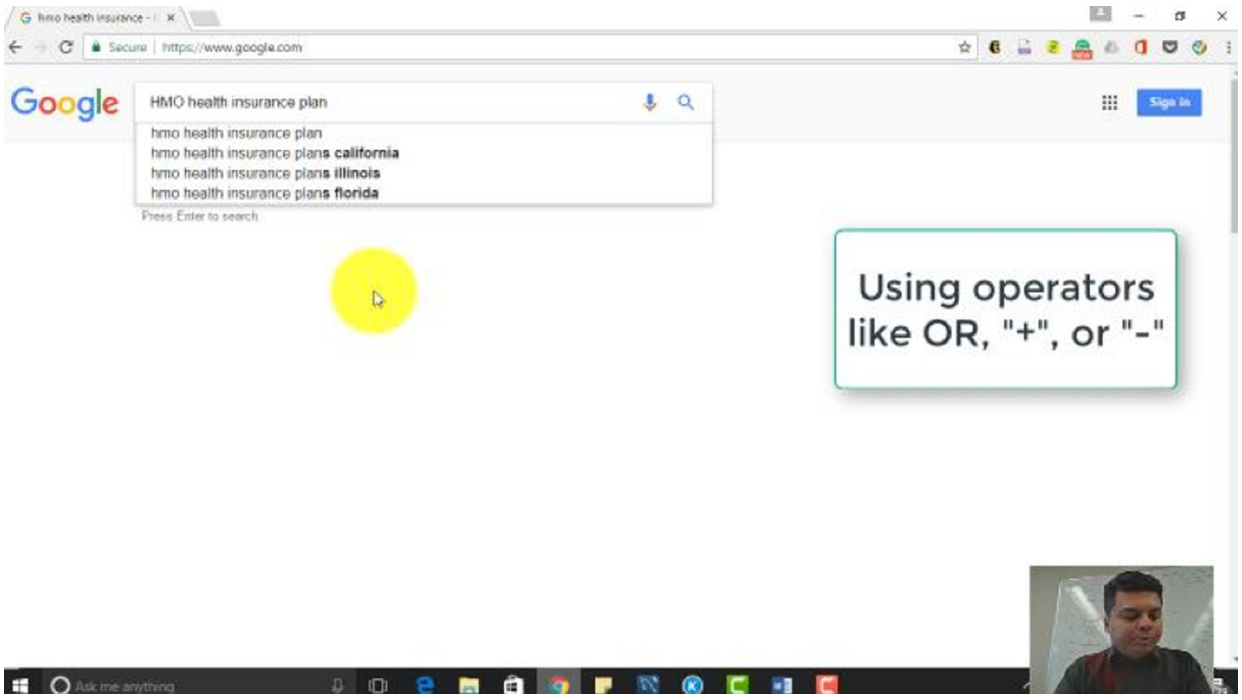

Fig. 2. Video screenshot of cognitive authority group.

environment and the participants were undergraduate students, the instruction from a professor can serve as an appropriate operationalization of cognitive authority. However, it is likely that the tips and advice from a university professor might be taken more seriously by the students than from an alternative type of cognitive authority (e.g., a search engine expert from Google, reference librarian) and thereby might increase the actual impact of advice. In the post-search interview, we asked participants if they knew the professor or had taken a class with him, and it turned out that no participants had directly interacted with the professor (e.g., taken a class, worked with the professor as advisee) before our study. Although there was no clear previous exposure to the influence of the professor's advice, biases from participants' life context (i.e., public university environment) might still affect the effects of search education. In this sense, the generalizability of the associated findings might be affected.

Participants in the peer-advice group watched a different 10-min video in which two students – one male and one female – were interviewed individually by an interviewer about how to search on Google. They also performed the same two search tasks on a computer screen. Again, participants were able to see both their faces and the screen on which they were searching and explaining their search strategies (thinking aloud).

Note that people in both videos described *exactly the same set of search tips using the same two tasks*. However, the person(s) (as peer advice or cognitive authority) who introduced the search tips and the language used to explain them varied between the two treatment conditions. The tasks used in the video were similar to the tasks assigned to participants in terms of task type and amount of effort required. Participants in the control group did not receive any treatment, nor did they know the existence of a treatment. Instead, they directly started the task section. Screenshots of treatment group videos are provided in Figs. 1 and 2, and the descriptions of the search tasks employed in the videos are provided as follows:

- Task 1 in the search education videos

  Suppose you are an employee of a university, you are trying to decide whether to choose a HMO health insurance plan or a PPO plan. Find resources that could help you decide which is best for your needs.

- Task 2 in the search education videos

  You want to work in the field of data science, but you don't have a background in this field. So you want to look for online courses that are offered to beginners with no programming and statistical knowledge. Find a course that may be suitable for your needs.

The main purpose of this study is to investigate the effect of cognitive authority and peer advice on individuals' search behavior. Therefore, the specific search tips introduced in each video were controlled but the people who presented the tips or the ways the content was delivered (i.e., sources of the search tips) were different. In the expert-advice video, the professor gave an online lecture covering how to search on Google, representing a cognitive authority on searching. In the student-advice video, two students

discussed search tips in the form of a casual interview guided by an interviewer, representing peer advice on searching. The search tips covered browsing and searching techniques such as using operators, formulating more specific and diverse queries, and spending more time on reading content pages. To illustrate how the search tips were introduced in different scenarios, we present some examples of video scripts as follows:

- Search tip 1: **Using operators**

  Excerpt 1: Expert advice video
  Professor: *So first, since I don't know much about HMO and PPO plans, I would like to start with looking for some general-level information about these two plans. In this case, using the relation word "OR" would be helpful because it allows me to get information about both of these plans by one single query. The operators like "OR", "+" and "-" can help you formulate longer and more accurate search queries and thus allows you to easily adjust the range of searching.*
  Excerpt 2: Student advice video
  Interviewer: *Hi Peter, could you try this search task, and tell us how you would proceed?*
  Student: *Hi, sure. So first, since I don't know much about these two plans, I would like to start with looking for some general-level information about these two plans. In this case, using the relation word "OR" would be helpful because it allows me to get information about both of these plans by one single query. The operators like "OR", "+" and "-" can help you formulate longer and more accurate search queries and thus allows you to adjust the range of searching.*

- Search tip 2: **Staying longer on content pages**

  Excerpt 1: Expert advice video
  Professor: *I always keep a tip in mind: use more time on reading these results than you think you should, don't just quickly scan or skip them. You will definitely benefit from these results if you do so.*
  Excerpt 2: Student advice video
  Interviewer: *How much time do you spend on each page?*
  Student: *Probably a few minutes. Like I mentioned before, I would spend quite some time on each page. Just to make sure that I don't miss anything important.*

Participants started the search tasks after watching the videos. The lab session was also different from the field sessions in two other aspects. First, each participant was only given up to 15 min to finish each task and they could decide to move on to the next task at any time within that 15 min. This time limit is only for the searching part, meaning that they were allowed to continue entering their answers after 15 min if necessary, though the large majority of participants could finish everything on time. Second, participants did not have to describe their search experience in a diary because they would be interviewed by one researcher regarding their experiences and thoughts immediately after the lab session when they still held the fresh memories of their experiences. Only one text box was provided for them to record the answers, though a few of them carried over the habit of writing down search experiences from the field session. The purpose of the lab study was to ensure that participants can start searching immediately after the treatment (i.e. search tips video) so that the immediate behavioral effects of search education can be properly captured.

Before each participant left, a semi-structured individual follow-up interview was conducted to further inquire about participants' experiences. The interview, as a complement to the diary and log data, was used to gain insights into participants' experiences with each search task in the first field session and the lab session as well as their opinions about the treatment. The researcher in charge of a session reviewed participants' diary responses and task answers of the first field session and prepared open-ended interview questions in advance. The researcher also took notes while watching a participant searching during a lab session. Based on the notes, a researcher constructed a list of questions tailored to each participant before going into a follow-up interview. Questions were primarily about search strategies and difficulties in finishing the tasks, and the videos. For example, some participants discussed why they considered specific tip useful or not useful. Participants also offered suggestions for improving the video treatments, which may be applied in future instructional design.

4.2.3. Field session 2

We started the second field session for each participant (by granting the access to the field session tasks) after he or she finished the lab session. Same as the first field session, participants were offered up to three days to finish two search tasks in naturalistic environment(s) of their choices. The second field session was designed to track any short-term or lasting effects of the search education received in the lab session. In both field sessions, participants received an email reminder reinforcing the deadline if they had not finished the tasks one day before the session ended. All user search behavioral data were automatically saved in a remote server by the Chrome browser plugin.

All participants were asked to answer two questions about their topic knowledge and search experience after finishing each task regardless of session or group: (1) How difficult was it to find the information you needed for this task? (*Perceived task difficulty*) (2) How knowledgeable were you on this topic? (*Topic knowledge*). Specifically, they were asked to evaluate their topic knowledge and perceived task difficulty using a 7-point Likert Scale.

4.3. Tasks

Participants finished two simulated search tasks in each session. All tasks were factual tasks with different end goals according to Li and Belkin (2008)'s definition (i.e., tasks locating facts or data in information systems). The first task in each session provided a specific goal and had one clear, correct answer, while the second task in each session had an amorphous goal and the answer was shaped by participants' consideration and interpretation of the factual information retrieved. The six tasks are:

- Field 1A (Factual specific)

  You saw a clip of a TV movie on a bus and want to watch the full movie. You don't know the name of the movie. You only saw that the heroine is an amateur sleuth who owns a bakery. In the clip you were watching, she was worried about her competitor stealing her business. You really like the actor who plays the heroine's friend, a dentist. Find out the name of that actor.

- Field 1B (Factual amorphous)

  Suppose you are an employee at a US non-profit organization. You are trying to decide whether to contribute money to your employer-offered 403(b) retirement plan or to a personal Individual Retirement Account. Find resources that could help you make the decision.

- Lab A (Factual specific)

  You are looking for a book titled Suede to Rest *by Diane Vallere that is not owned by Rutgers University Library. You want to find the closest library that owns the book and how you can check it out.*

- Lab B (Factual amorphous)

  You are visiting Beijing and want to find information about Beijing Capital Airport (PEK). Find information to answer the following questions: what shuttles and public transportation connect the airport to downtown; what hotels are close to the airport; what rental car services do they have; and where is cheap parking around the airport.

- Field 2A (Factual specific)

  Your friend really likes a pair of shoes [the picture of the shoes was shown to the participants] that he saw worn by the hero in the South Korean TV series Pinocchio. You want to buy these shoes as his birthday gift, but the TV series did not specify the brand. Find out the brand and style of the shoes.

- Field 2B (Factual amorphous)

  You are applying for Ph.D. programs in the area of Library and Information Science (LIS). You want to find out what are the best LIS programs in the US and which of them guarantees at least four years of funding for Ph.D. students. You also want to consider their job placement within the last three years. Did their graduates find tenure-track positions? Taking these factors into consideration, find five LIS programs that you want to apply to, and provide your reasons.

Among the six search tasks, Field 1B and Lab B were adapted from Text Retrieval Conference (TREC) 2014 session track tasks,[1] while the rest were designed by us.

4.4. Pilot test

This study used a convenience sampling to recruit pilot test participants by sending out email invitations. Five pilot test participants finished the study. This helped us verify if (1) the study system was functioning well, (2) the tasks and survey questions could be understood with no confusion, and (3) the tasks and time limit were reasonable (not too difficult or too easy to be completed in 15 min). Pilot test participants were interviewed in person about their suggestions for the study procedure (e.g., if there was any confusion about the tasks and questionnaires; if the time was enough). Minor modifications were made to the wording of some questions, tips and tasks based on the feedbacks from the pilot study participants.

4.5. Behavioral measurements

To examine the behavioral effects of peer advice and cognitive authority, this study employed six behavior measurements that are directly associated with the search tips introduced in the instructional videos, and examined the differences in users' Web search

[1] https://trec.nist.gov/data/session/2014/topictext-890.txt.

Table 1
Search tips, expected changes in behavior, and the associated segments from video scripts.

Search Tip	Expected change in the treatment groups	Associated segment from the instructional video scripts
Use search operators in queries, such as +, -, OR, and quotation marks (Moraveji et al., 2011)	**H1**: Number of search operators used↑	The operators like "OR", "+" and "-" can help you formulate more accurate search queries and thus allows you to easily adjust the range of searching…I feel I'd prefer PPO, but I want to learn more about the advantages…so I type in "advantages of PPO". Besides, I also put the phrase in a quote [type in "advantages of PPO"] …
Use more diverse terms and formulate more unique queries	**H2**: Number of unique queries↑ **H3**: Number of unique query terms↑	… Also, using more diverse terms in queries and formulating unique, original queries can help broaden the range of search and may increase the possibility of finding useful information…
View more content pages/search results	**H4**: Number of content pages viewed↑	[Click on the results on second search result page] Choosing a health insurance plan is a pretty complicated task so I normally go through the first result page and read the results one by one, I may even check the results and webpages on the second and third pages if necessary… And I would probably go through most of the results, rather than just looking at the top ones
Spend more time on reading search engine result pages (SERPs) and content pages	**H5**: Mean dwell time on SERP↑ **H6**: Mean dwell time on content page↑	For the results on the first result page, I always keep a tip in mind: use more time on reading these results than you think you should, don't just quickly scan or skip them. You will definitely benefit from these results if you do so… I don't know exactly what I'm looking for, so I would keep reminding myself to spend more time on reading the result lists as well as each result page. I may miss something if I move too quickly…

Note: H stands for "hypothesis". In the "associated segments from video scripts" column, sentences in brackets describe the corresponding search actions taken by the professor and peer advisor when they introduce the associated search tips in instructional videos.

behaviors across different tasks and groups. Table 1 clarifies the connections between search tips, expected changes in the corresponding behavioral measurements (hypotheses), and the associated segments from the instructional video scripts.

Among the four search tips, the first one introduced in the Table 1 was adapted from Moraveji et al. (2011), while the other three search instructions were developed based on the discussions among the researchers of this study. It is worth noting that in the original search tips Moraveji et al. (2011) were tailored for the predefined task contexts and were designed to improve the search performance. However, our study focused on the effects of the ways that the search tips or persuasions were presented (instead of the specific content of search tips) on Web search behavior (instead of search performance or outcome). With respect to the hypotheses, our general expectation was that participants become more active in multiple aspects of search activities (e.g., querying, browsing) under the influence of the designed search tips. Note that we did not make any hypothesis regarding the potential difference between the effect of peer advice and that of cognitive authority. This is because there was no previous literature that directly compares the behavioral effects of these two factors in user education or persuasion context and thus we need to gain the knowledge from our analysis.

The measurements defined in this study can be classified into two categories. The first category is query behavior, including the frequency of usage of search operators (+, −, OR, AND, and quotation marks), number of unique queries used within a task session (repeated queries were only counted once), and number of unique query terms used in formulating queries. The other category is page viewing behavior. We first calculated the number of content pages viewed by participants within each task session. In terms of dwell time features, for the dwell time on SERP measure, we used the difference between the timestamp of starting a SERP and the timestamp of leaving a SERP to calculate SERP dwell time. The mean dwell time on SERP was the average SERP dwell time per query within each task session. Similarly, we used the timestamp of clicking a content page (result) and the timestamp of leaving a content page to compute content page dwell time and calculated the average dwell time on each content page within each task session. To reduce the noise from distraction and digression in field sessions, we discussed and manually removed some of the URLs which are clearly irrelevant to our task (e.g., personal Facebook page, personal Twitter page, Gmail page, Rutgers Sakai page).

4.6. Analysis

The first field session served as the pretest for subsequent treatment effect analysis. We statistically tested the differences in search behavior and task perception across different groups in the first session, aiming to check if users in different groups started from similar situation. Then, to measure how users' search behaviors evolved under the influence of peer advice and cognitive authority respectively, we tested the differences in users' search behaviors across different treatment situations (control, student advice, expert advice) in the lab session (RQ1 and RQ2). We also compared the search behaviors of users from two treatment groups to that of the control group in the second field session, aiming to investigate if the behavioral effects lasted in a short period of time when the explicit search tips and instructions were removed (RQ3).

Besides the quantitative data collected by browser plugin, all post-search semi-structured individual interviews were transcribed for qualitative data analysis. Since this paper only reports the effects of treatment videos, the discussion related to participants' understanding and usage of the tips introduced in the videos were separated from the rest of the data and analyzed in

NVivo.[2] One person in the research group first coded two interviews by drawing codes inductively from the data, after which, all group members met and reviewed the coding to minimize personal bias. Then the coder finished the rest of the interviews. As only a limited amount of codes was generated and the coding process was straightforward, no inter-coder reliability was calculated.

5. Results

This study collected search behavior and diary data from 185 search task sessions.[3] To protect participants' privacy and clean out clearly divergent pages, we manually removed potentially sensitive information and the associated Web pages collected in the field sessions (e.g., emails, personal Facebook pages, irrelevant YouTube videos, Amazon pages, other personal account pages) before conducting statistical analysis.

To explore the immediate behavioral effects of peer advice and cognitive authority (RQ1 and RQ2), this paper examined the differences in search behaviors across three different groups/treatment situations (i.e., control group, peer advice group, cognitive authority/expert advice group) based on lab session data. To explain the potential short-term effects of the treatments (RQ3), we examined the behavioral differences across the three groups in the second field session. The results of statistical analysis are provided in the following sections.

5.1. Pre-treatment baseline condition: Cross-group comparison in field session 1

Before analyzing the effects of search advice, we examined the differences in search behavior, topic knowledge, and perceived task difficulty across groups in the baseline condition before participants received any treatments (i.e., field session 1), aiming to test if participants started from similar situations. Since none of the behavioral dataset from the three conditions was normally distributed, this study employed Kruskal-Wallis test for cross-group comparison. In addition, we conducted statistical analyses under the two types of search tasks (i.e. factual specific and factual amorphous) respectively in order to the further explore the behavioral effects of user education of different types in different task contexts.

According to the results of Kruskal-Wallis tests presented in Table 2, no statistically significant behavioral differences was found among the three groups in the first field session. Similarly, according to the results presented in the first column of Table 5, there was no significant difference in topic knowledge or perceived task difficulty among groups. These results indicate that participants from different groups did not differ significantly in search behavior and task aspects when search advice was not introduced. Based on this finding, it is reasonable to assume that participants in different groups started from the same situation before the search tips were introduced. The following sections seek to illustrate how different aspects of users' search behavior varied under the influences of search advice after departing from the same situation.

5.2. Cross-group comparison: Lab session

This paper statistically examined the differences in search behavior among the three groups in both lab and second field session, aiming to understand the impacts of different sources of search advice. Since none of the search behavior dataset in different conditions was normally distributed, Kruskal-Wallis test was employed for cross-group statistical analysis in this section. In the cases where the differences among groups were statistically significant, we continued our analyses with Dunn's post-hoc pairwise test in order to examine the between-group differences. Given that users often behave differently in different task contexts (Jiang et al., 2014; Li & Belkin, 2008), we investigated the behavioral effects of search advice in the two search task contexts separately.

As the response to RQ1 and RQ2, the results of the lab session cross-group Kruskal-Wallis tests and the associated post-hoc pairwise tests are presented in Table 3. Since this study employed nonparametric tests for the comparative analysis, Table 3 reports the medians and inter-quartile ranges (IQR) (instead of means and standard deviations) of the search behavior measures. In summary, the results demonstrate that (1) both peer-advice- and expert-advice-based (cognitive authority) user education generated significant impacts on multiple aspects of participants' Web search behaviors (e.g., usage of search operators, content-page-viewing behavior); (2) according to the results of post-hoc pairwise tests, in most cases, there was no significant difference between peer advice and cognitive authority in terms of the immediate behavioral impacts; (3) In general, the user education or instructional intervention designed here created relatively broader behavioral impacts in factual amorphous tasks than in factual specific tasks. Specifically, for example, the forth search tip (spending more time on examining SERPs and content pages) operationalized as video-based user education led to the significant increases in mean dwell time on SERP and content page in factual amorphous task only. This may be because the relevance of the visited web pages and results was relatively clear and straightforward in the factual tasks with clearly-defined, specific goals. Therefore, despite the potential effects of user education, participants might not feel necessary to actually follow these instructions (i.e., spending extra time on SERPs and content pages) when working on factual specific tasks. However, it is worth noting that although the second search tip introduced in Table 1 sought to encourage participants to formulate more unique queries, it might be difficult for participants to come up with extra useful queries in tasks which do not have a clear goal. Thus, no statistically significant effect on the total amount of unique queries was found in factual amorphous task.

[2] https://www.qsrinternational.com/nvivo/home.

[3] Because of system errors, one search task sessions and 122 SERPs were lost in data collection. Irrelevant webpages (e.g., Gmail, student portal) were excluded from data analysis.

Table 2
Cross-group comparison (baseline/field session 1): Median (IQR).

Behavior and task context	Control	Peer advice	Cognitive authority	Kruskal-Wallis test
Number of search operators (FS)	0(0)	0(1)	0(0)	$\chi^2 = 0.68$, $p = .71$
Number of search operators (FA)	0(1)	1(2)	0(0)	$\chi^2 = 5.28$, $p = .07$
Number of unique queries (FS)	4(3)	5(4)	3.5(5.5)	$\chi^2 = 1.35$, $p = .51$
Number of unique queries (FA)	2(7)	2(1)	3(2)	$\chi^2 = 3.12$, $p = .21$
Number of unique terms (FS)	15(11)	16(19)	12.5(15)	$\chi^2 = 0.67$, $p = .71$
Number of unique terms (FA)	15(18)	10(4)	9.5(3)	$\chi^2 = 1.99$, $p = .37$
Number of content pages (FS)	7(6)	10(22)	12(17)	$\chi^2 = 3.80$, $p = .15$
Number of content pages (FA)	5(19)	5(3)	6(5)	$\chi^2 = 3.04$, $p = .22$
Mean dwell time SERP (FS)	22.2(8.23)	22.9(41.9)	11.87(27.13)	$\chi^2 = 1.83$, $p = .40$
Mean dwell time SERP (FA)	23.22(18)	10.54(18.5)	15.4(32.52)	$\chi^2 = 1.23$, $p = .54$
Mean dwell time content (FS)	33.71(10.03)	37.29(47.7)	25.99(47.14)	$\chi^2 = 0.07$, $p = .97$
Mean dwell time content (FA)	61.33(73.92)	73.57(95.36)	114.13(51.5)	$\chi^2 = 3.33$, $p = .19$

Note: FS: factual specific task. FA: factual amorphous task.

Table 3
Cross-group comparison (lab session): Median (IQR).

Behavior and task context	Control	Peer advice	Cognitive authority	Kruskal-Wallis test and Dunn's post-hoc pairwise test
Number of search operators (FS)	0(0)	1.5(3)	1(3)	**$\chi^2 = 9.04$, $p = .01$. P > C**, E > C***, P vs E (n.s.)
Number of search operators (FA)	0(1)	3(2)	4.5(3)	**$\chi^2 = 21.90$, $p = .00$. P > C**, E > C****, P vs E (n.s.)
Number of unique queries (FS)	4.5(3)	5(7)	9(8)	**$\chi^2 = 8.05$, $p = .02$. P > C**, E > C****, P vs E (n.s.)
Number of unique queries (FA)	6(2)	7(3)	8(6)	$\chi^2 = 1.63$, $p = .44$.
Number of unique terms (FS)	13(7)	25(18)	26(13)	$\chi^2 = 5.47$, $p = .06$.
Number of unique terms (FA)	12(2)	19(4)	21.5(14)	**$\chi^2 = 20.80$, $p = .00$. P > C**, E > C****, P vs E (n.s.)
Number of content pages (FS)	14(6)	43(38)	31.5(4)	**$\chi^2 = 18.44$, $p = .00$. P > C**, E > C**, P > E****
Number of content pages (FA)	15(5)	28(12)	32(11)	**$\chi^2 = 21.12$, $p = .00$. P > C**, E > C****, P vs E (n.s.)
Mean dwell time SERP (FS)	7.7(12.18)	12.05(15.36)	10.99(4.49)	$\chi^2 = 1.75$, $p = .42$.
Mean dwell time SERP (FA)	6.71(1.75)	16.61(3.63)	14.45(5)	**$\chi^2 = 16.78$, $p = .00$. P > C**, E > C****, P vs E (n.s.)
Mean dwell time content (FS)	11.56(6.98)	16.38(8.2)	18.11(12.4)	$\chi^2 = 3.55$, $p = .17$.
Mean dwell time content (FA)	13.66(6.55)	25.36(3.89)	28.6(12)	**$\chi^2 = 21.62$, $p = .00$. P > C**, E > C***, P vs E (n.s.)

Note: FS: factual specific task. FA: factual amorphous task. C: control group, P: peer advice group, E: expert advice/cognitive authority group. * < .05, ** < .01. n.s.: no statistically significant difference.

In this section, we investigate the behavioral impacts of user education and compare the effects of instruction interventions of different types. The following section takes a step further and compares the short-term influences of peer advice and cognitive authority on Web search behavior when the explicit search interventions were removed. To this end, we present the results from the cross-group comparative analysis based on the data collected from the second field session.

5.3. Cross-group comparison: Field session 2

To answer the RQ3, we tested the short-term behavioral impacts of peer advice and cognitive authority in naturalistic settings where the explicit interventions were removed. The results are presented in Table 4. In summary, we found that (1) The short-term

Table 4
Cross-group comparison (field session 2): Median (IQR).

Behavior and task context	Control	Peer advice	Cognitive authority	Kruskal-Wallis test and Dunn's post-hoc pairwise test
Number of search operators (FS)	0(0)	0(0.5)	4(2)	**$\chi^2 = 17.53$, $p = .00$**. P vs C (n.s.), **E > C**, E > P****
Number of search operators (FA)	1(2)	1(4)	6(5)	**$\chi^2 = 13.59$, $p = .00$**. P vs C (n.s.), **E > C**, E > P****
Number of unique queries (FS)	4(3)	6.5(10)	5.5(2)	$\chi^2 = 1.44$, $p = .49$.
Number of unique queries (FA)	5.5(11)	4(4)	8(22)	$\chi^2 = 4.38$, $p = .11$.
Number of unique terms (FS)	13(9)	12(8)	20(6)	**$\chi^2 = 11.15$, $p = .00$.** P vs C (n.s.), **E > C**, E > P****
Number of unique terms (FA)	16(11)	11(2)	25(11)	**$\chi^2 = 13.98$, $p = .00$.** P vs C (n.s.), **E > C**, E > P****
Number of content pages (FS)	12(19)	14.5(15)	22.5(15)	$\chi^2 = 4.27$, $p = .12$.
Number of content pages (FA)	16(13)	27(20)	27(6)	$\chi^2 = 5.43$, $p = .07$.
Mean dwell time SERP (FS)	24.17(31.37)	15.29(14.04)	9.43(14.2)	$\chi^2 = 4.21$, $p = .12$.
Mean dwell time SERP (FA)	10.83(12.13)	17.63(12.33)	10.33(8.4)	$\chi^2 = 1.78$, $p = .41$.
Mean dwell time content (FS)	15.74(21.63)	22.38(9.18)	24.62(9.15)	$\chi^2 = 1.88$, $p = .39$.
Mean dwell time content (FA)	29.27(23.44)	26.70(25.28)	31.29(20.28)	$\chi^2 = 0.74$, $p = .69$.

Note: FS: factual specific task. FA: factual amorphous task. C: control group, P: peer advice group, E: expert advice/cognitive authority group. *: < 0.05, **: < 0.01. n.s.: no statistically significant difference.

Table 5
Topic knowledge and perceived task difficulty across groups: mean (S.D.).

Measures	Field session 1	Lab session	Field session 2
Topic knowledge (Factual specific)	C: 3.58(2.19), P: 3(1.49), E: 3.5(2.44) $\chi^2 = 2.00$ $p = .37$	C: 3.36(1.57), P: 3.7(1.16), E: 3.4(1.17) $\chi^2 = 1.11$, $p = .57$	C: 2.64(1.4), P: 2.7(2.21), E: 2.7(1.23) $\chi^2 = 3.39$, $p = .18$
Topic knowledge (Factual amorphous)	C: 2.42(1.16), P: 2.6(0.97), E: 2(1.49) $\chi^2 = 1.27$, $p = .53$	C: 3.36(1.12), P: 3(2.58), E: 3.1(1.73) $\chi^2 = 0.81$, $p = .67$	C: 3.3(1.34), P: 2.9(1.37), E: 2.6(1.07) $\chi^2 = 0.58$, $p = .75$
Perceived difficulty (Factual specific)	C: 3.42(1.67), P: 3.8(1.69), E: 1.63(1.19) $\chi^2 = 1.05$, $p = .59$	C: 2.73(1.34), P: 4.2(1.62), E: 3.3(2) $\chi^2 = 1.42$, $p = .49$	C: 4(2.37), P: 4.1(2.13), E: 2.8(1.39) $\chi^2 = 2.43$, $p = .30$
Perceived difficulty (Factual amorphous)	C: 3(1.86), P: 4.4(1.84), E: 4.1(0.99) $\chi^2 = 4.16$, $p = .13$	C: 3.55(1.51), P:3.9(2.02), E: 4.3(1.34) $\chi^2 = 1.64$, $p = .44$	C: 4.3(1.42), P: 5.2(1.81), E: 5.3(1.64) $\chi^2 = 0.51$, $p = .77$

Note: C: control group. P: peer advice group. E: expert advice (cognitive authority) group.

impacts were retained only in cognitive authority group. No significant difference was found between the control group and peer advice group in terms of the values of the six behavioral measures. (2) The short-term behavioral impacts of cognitive authority only appeared in querying behavior (specifically, in the usage of search operators and unique query terms). No significant impact was found on content-page-viewing and dwell time measures. These results indicate that (1) compared to peer advice, cognitive authority (operationalized as expert advice) is more likely to generate the behavioral effects that can last at least in a short period of time. (2) Explicit search education or instruction is more likely to cause short-term continuing effects on query formulation strategies than on page-viewing and browsing behaviors. In other words, it may be relatively difficult to change users' established strategies in page-viewing and examination via search instruction.

5.4. Topic knowledge and perceived task difficulty

To answer RQ1, 2, and 3, previous sections present the analyses of a series of cross-group behavioral variances in both lab (testing the immediate effects, RQ1 and RQ2) and second field sessions (testing the short-term effects, RQ3). However, it is still not clear whether the behavioral variations across groups resulted from any significant differences in participants' topic knowledge and perceived task difficulty. To addressing this issue, we statistically tested the differences in topic knowledge and perceived task difficulty across groups in the three sessions respectively. We first ran Shapiro-Wilk tests to test the normality of data distribution for each variable in each condition in order to determine which statistical model we should use. The results showed that the answer data of every variable was normally distributed in all sessions. Thus, we employed one-way ANOVA to test the cross-group differences in topic knowledge and perceived task difficulty. The results are presented in Table 5.

According to the results in the second and third columns in Table 5, no significant cross-group difference in topic knowledge or perceived task difficulty was found. These results suggest that the observed search behavior variations cannot be simply explained by the differences in participants' topic knowledge and perceived difficulty across groups, and that the video-based user education did not lead to any significant increase in users' perceived task difficulty. These results can help further clarify the behavioral effects of search advice in different sessions and treatment groups and support our arguments.

5.5. Users' voice: Evidences from the post-search semi-structured individual interviews

Although the log data collected by browser plugin was able to show the extent to which the introduced search tips affected users' search behaviors, it did not reveal participants' opinions and attitudes towards the video content and the underlying reasons that motivated them to (or not to) adopt those tips. We were able to gain further understanding of the impact of treatment videos on participants' search behaviors from the semi-structured post-search interviews arranged at the end of lab sessions.

Both the content and the presentation of the instructional videos were discussed in the post-search interviews. To encourage participants to reflect on their search experiences and usage of search advice, the researcher in charge of a lab session always started with a general question regarding the video (i.e., What did you think of the video?) and let the participant decide which aspects and/or tips to discuss. Among all the tips presented, using operators and quotation marks was the most frequently discussed tip in the interviews. Participants found this tip useful and easy to memorize, like P18 pointed out: "*A lot of the time when I'm searching for something and I type out a whole phrase, it would scatter it, or they would just use one part of the phrase and that's not exactly what I wanted. But the quotations really do help.*" Interestingly, while a few participants have learned or heard of this tip before, they rarely tried those on a search engine or they did not know how to use them: "*With the plus and minus I didn't realize that you have to leave no space afterwards. I've heard of using that before, but I didn't know you couldn't put a space after the plus or the minus so that was good tip*" (P05). These qualitative evidences support the finding that the statistically significant cross-group differences in the usage of search operators were found in both lab session (immediate effect) and second field session (short-term effect).

This study focuses on the effects of cognitive authority and peer advice on search behaviors rather than on search performance, and thus the tips and advice were not designed to effectively improve users' search outcomes (e.g., chances of finding better results). However, participants sometimes still felt that they succeeded because of those tips and hence tended to employ the search tips when

they encountered obstacles in search. For instance, P27 brought up that "*I used Boolean a lot, particularly 'AND'. This would have sifted through a lot more unnecessary information if I had not used that.*" As a result, she "*attribute(s) most of that success to using my BOOLEAN knowledge.*" There were also times when the participants did not use the tips at first but tried them when struggling with finding information. This suggests that users may not think that they would need some advanced search techniques until they really have difficulties in searching. This finding helps explain the finding that when comparing the number of search operators used across different groups in the lab session, there was a larger gap between control group and treatment groups (especially cognitive authority group) in factual amorphous task than that of factual specific task. A possible reason is that participants might struggle with finding useful webpages more often in factual amorphous task and thus they were more likely to be motivated by the instructional video and thereby to be open to using more search operators and quotation marks in query formulation. In other words, this result demonstrate that participants were more likely to act on the search advice in goal-amorphous, intellectually challenging tasks, even if some of the advice (e.g., operator "AND") has no objective influence on search performance.

Other than using operators and quotation marks, the treatment videos also introduced other tips such as viewing more content pages and spending more time on reading content pages. However, those tips were not equally appreciated by the participants as the former ones. P21 brought up that the 15-min time constraint prevented her from looking at more pages: "*If I had more time I would have looked through everything properly without worrying about running out of time.*" This finding echoes the corresponding analysis results that no significant cross-group difference was found in mean dwell time on content pages in factual specific tasks. In addition to the explanation given by the P21, this may also be because using operators and quotation marks as a search tip is more salient and noticeable than the other tips which are mainly concerned with common search behavioral features (e.g., number of pages to visit, dwell time on content pages). Another possible reason is that users can determine the relevance of a content page and find useful material in a quick manner (e.g., in 30 s or less) and therefore may not need to significantly increase their content page reading time when performing factual search tasks.

Participants also offered suggestions that may be considered and implemented in future user education or intervention design, particularly regarding the presentation of the search advice. For example, P30 suggested that we can include a summary of tips at the end of a video: "*Maybe at the end having a summary of things that you can do. A lot of times you said something and then you might zone out, instead of going back you just have a summary*." We did consider this when shooting the videos. Thus, the main tips were repeated at least once in the videos in case participants miss or forget anything. However, several participants considered the repetitions unnecessary and hoped to shorten the video, "*I felt like the video was a little long. I think they went over similar things, so after watching the first section and then watching the second section, it was a little bit repetitive*" (P14). Thus, having a search tip summary that participants can choose to review or skip could be a viable replacement of repeating the same content in search instruction.

6. Discussion

Table 6 offers a concise summary of the findings from cross-group analyses and identifies the hypotheses (introduced in the second column of Table 1) that are empirically confirmed by the results. Overall, based on the analysis of 185 task session data generated by 31 participants in two field and one lab sessions, this study demonstrates that peer advice and cognitive authority operationalized as video-based search advice can cause both immediate (RQ1 and RQ2) and short-term (RQ3) effects on certain aspects of users' Web search behaviors in different task contexts. The related evidences extend the classical arguments and theories concerning the behavioral impacts of peer advice and cognitive authority (e.g., Cohn et al., 2014; Gardner & Steinberg, 2005; Rieh, 2002; Wilson, 1983) to the context of Web search, and thereby reveals the potential value of leveraging these effects in search engine user education and instructional intervention design.

As the response to RQ1 and RQ2, according to the findings from the lab-session-based analysis, we found that (1) overall, both types of treatments or instructions generated significant impacts on multiple aspects of participants' Web search behaviors (e.g., usage of search operators, page-viewing behavior); (2) according to the results of post-hoc pairwise tests, in most cases, there was no significant difference between peer advice and cognitive authority in terms of the immediate behavioral impacts; (3) In general, the user education designed here had relatively broader behavioral impacts in factual amorphous tasks than in factual specific tasks. Note that the effects of search advice on querying (i.e. usage of search operators, usage of unique queries or unique terms) and number of content page visited are statistically significant in both types of tasks, indicating that it is possible to shape these aspects of search behavior in tasks of different types. Previous literature has pointed out that users may not even have a reasonable query to begin with, especially in exploratory tasks (Hendahewa & Shah, 2017), and thus they may need help with formulating effective queries (Aula, Khan, & Guan, 2010). These findings can partially explain why participants were more willing to accept and apply the search advice related to query formulation strategies.

Table 6
The hypotheses (H1-6 in Table 1) confirmed in different task-group-session combinations.

Task context	Treatment group	Lab session	Field session 2
Factual specific	Peer advice	H1, H2, H4.	
	Cognitive authority	H1, H2, H4.	H1, H3.
Factual amorphous	Peer advice	H1, H3, H4, H5, H6.	
	Cognitive authority	H1, H3, H4, H5, H6.	H1, H3.

However, not all search tips were effective in both task types. The forth search tip (spending more time on examining SERPs and content pages) led to the significant increases in mean dwell time on SERP (H5) and content page (H6) in factual amorphous tasks only. A possible reason is that the relevance of the visited web pages and results was relatively clear and straightforward in the factual tasks with clearly-defined, specific goals. Therefore, despite the behavioral effects of search advice, participants might not feel necessary to actually follow these instructions and to spend more time on assessing SERPs and Web pages when working on factual specific tasks. Consequently, we did not find any significant cross-group difference in the dwell time on SERPs and content pages in the context of factual specific task.

With respect to the RQ3, in contrast with the finding that peer and cognitive authority were equally effective in generating immediate behavioral impacts, the short-term effects of search advice on query formulation (usage of search operators, number of unique terms used) only existed in cognitive authority groups. Specifically, the results indicate that participants from the cognitive authority group used significantly more search operators (H1) and unique terms (H3) in the second field session where the explicit search advice were removed, and that no significant behavioral difference was found between the control group and peer advice group in terms of the values of the six behavioral measures. This result suggests that compared to peer advice, cognitive authority is more likely to generate lasting effects on Web search behavior. It also proves that in addition to the content features of search advice and tips (cf. Moraveji et al., 2011; Yamamoto & Yamamoto, 2018), the difference in the source of user search education (i.e., from peer or cognitive authority) may also lead to divergent behavioral effects on some components of Web search. In this sense, this study at least partially reveals the underlying complexity and multidimensionality of user education in the context of Web search.

In addition, we also found that the search advice did not cause any significant short-term impact on the content-page-viewing and dwell time measures. This result suggests that the externally assigned search instruction or advice are more likely to cause short-term continuing effects on query formulation strategies than on page-viewing and browsing behaviors. In other words, for search intervention designers and researchers, it may be difficult to shape users' established strategies and habits in page viewing and result examination via explicit search instruction.

As always, there are several limitations to this study as well as needs for future research efforts. First, peer advice and cognitive authority were operationalized as two treatment videos (student's advice and expert/professor's advice) given to participants. There may be other intervention techniques that can be considered and implemented in presenting search advice in future research. In addition, since this study was conducted within a university environment and the participants were undergraduate students, instructions from a professor might be taken more seriously than from another type of cognitive authority (e.g., a search engine expert from Google) and thereby might increase the impact of advice. Although there was no clear previous exposure to the influence of the professor's advice, biases from participants' life context might still affect the actual effects of search education. Hence, it is not clear if the conclusions drawn from our results can be generalized to other populations and contexts (e.g., employees in tech companies, lawyers, library users). In future studies, comparing different types of cognitive authorities in different study contexts (e.g., university-based academic research, online search engine A/B testing) may further enhance our understanding of the behavioral impacts of Web search education. In terms of the instruction language, the differences in the language used to describe search tips might also result in immediate and short-term behavioral variations across groups. To minimize this potential effect, we used same or very similar terms [showing on the screen] for same search tips in peer advice and cognitive authority groups (see Figs. 1 and 2). Overall, as an initial attempt, this study reveals the value of employing video-based instructional interventions in Web search education and taking into account the factors from individuals' social contexts in user education design.

Second, the search tips were designed based on previous IIR studies (e.g., Moraveji et al., 2011), the results of a pilot study, as well as the discussions among the researchers of this study. Some of the introduced search advice was outdated (e.g., AND) and was no longer supported by the current search engines, and a series of current Google search operators (e.g., $, #, @) were not included in our search education video scripts. However, it is worth noting that since the focuses of this study are the sources of search intervention or the way in which we present the search tips (i.e., as peer advice versus as cognitive authority) rather than the specific contents of search tips or actual search outcomes, the conclusions drawn from our analysis can still contribute to the understanding of the potential value of different types of advice-based search education. Nevertheless, to design more useful, applicable Web-based search interventions, future research should include more effective, up-to-date operators and tips in search education, especially when the goal of research is to evaluate the positive effects of Web-search education on search outcomes and experiences.

Third, we did not show any video in the control group. Participants in the control group directly started their search sessions based on the task requirements. Thus, some may argue that the exposure to a search intervention video actually triggered a desire to do well or work hard for the advice givers rather than motivating participants to follow the search tips. However, it is worth noting that to encourage participants to take our study seriously, we told participants at the beginning that the top five performers judged by us would receive a $20 bonus. The performance was judged based on how detailed and organized their answers were (e.g., the quality of their answers), and all participants were informed of this criterion after they registered for the study. We believe that the extra bonus here already triggered a desire to do well not only among the advice receivers, but also among the participants in the control group. Thus, we think the difference in search behavior between control group and treatment groups can at least to some extent reveal the behavioral effects of the search tips of different types. Nevertheless, for a more rigorous study design, in the control group we need a video asking the participants to please work hard and remind them of the bonus structure, but give no clear advice on how to obtain success. Adding this "placebo" video may help us better separate the potential effect of the exposure to a search intervention video from the actual behavioral effect of the search tips and interventions from different sources.

Fourth, participants' level of engagement in searching gradually dropped as they were getting closer to the end of the study. This is mostly reflected in the decreasing length of diary entries. This may also be a reason that led to insignificant effects of the treatment on some search behaviors in the second field session. Given this natural tendency, designing strategies to maintain participants'

engagements during the entire study is critical in behavioral studies and thus should be taken into account in future research. Also, this study only examines the short-term lasting effects of cognitive authority and peer advice, more longitudinal studies could inform the understanding of the long-term learning effects of Web searching education.

And, finally, in order to control task-level variables, only factual tasks were employed this time. As it is shown in the findings, search education videos may have a bigger impact on exploratory search tasks. Therefore, examining the influence of instructional interventions in other types of tasks, particularly open-ended, exploratory search tasks that produce new ideas and designs (i.e., intellectual tasks), will be the next step of future research.

7. Conclusion

In contrast to the long-lasting attention to the algorithm and interface dimensions of search, there is a lack of research concerned with the potential effects of different types of user education on search behavior. To address this research gap, the work reported in this paper examined the immediate and short-term impacts of cognitive authority and peer advice on Web search behaviors in the context of search engine user education. Thirty-one participants completed a total of 185 search-task sessions (62 lab sessions, 123 field sessions) taking place in a controlled lab environment and the natural settings of their choices.

Overall, as the response to RQ1 and 2, this study found that both cognitive authority and peer advice (operationalized as search education videos) generated immediate impacts on several aspects of Web search behavior, and that the behavioral impact of search advice was broader in factual amorphous task than in factual specific task. As the response to RQ3, our results indicate that in the second field session, cognitive authority generated statistically significant impacts on the usage of search operators and unique terms in query formulation, whereas the behavioral effect of peer advice faded away when the explicit search tips were removed. These findings empirically demonstrate the behavioral impacts of peer advice and cognitive authority in user education and suggest that (1) peer advice and cognitive authority can be equally useful in motivating immediate behavioral changes; (2) framing search tips as the advice from cognitive authority is more likely to generate continuing, short-term effects on Web search behaviors. Future studies need to further extend this line of search-engine user education research to open-ended, cognitively-demanding, exploratory search tasks in order to further explore the potential of utilizing the information and guidance from cognitive authority and peers in search education.

Acknowledgments

The work reported here is supported by the Institute of Museum and Library Service (IMLS) grant no. LG-81-16-0025-16. The authors would like to thank the two anonymous referees of this paper for their insightful comments which helped to improve the manuscript significantly.

References

Aula, A., Khan, R. M., & Guan, Z. (2010). How does search behavior change as search becomes more difficult? *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 35–44). ACM.

Bates, M. J. (1979). Information search tactics. *Journal of the American Society for Information Science and Technology, 30*(4), 205–214.

Belkin, N. J., Marchetti, P. G., & Cool, C. (1993). BRAQUE: Design of an interface to support user interaction in information retrieval. *Information Processing & Management, 29*(3), 325–344.

Brajnik, G., Mizzaro, S., Tasso, C., & Venuti, F. (2002). Strategic help in user interfaces for information retrieval. *Journal of the Association for Information Science and Technology, 53*(5), 343–358.

Cao, H., Jiang, D., Pei, J., He, Q., Liao, Z., Chen, E., et al. (2008). Context-aware query suggestion by mining click-through and session data. *Proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 875–883). ACM.

Capra, R., Arguello, J., Crescenzi, A., & Vardell, E. (2015). Differences in the use of search assistance for tasks of varying complexity. *Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval* (pp. 23–32). ACM.

Cohn, A., Fehr, E., Herrmann, B., & Schneider, F. (2014). Social comparison and effort provision: Evidence from a field experiment. *Journal of the European Economic Association, 12*(4), 877–898.

Cucerzan, S., & White, R. W. (2007). Query suggestion based on user landing pages. *Proceedings of the 30th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 875–876). ACM.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117–140.

Frey, B. S., & Meier, S. (2004). Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment. *American Economic Review, 94*(5), 1717–1722.

Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. *Developmental Psychology, 41*(4), 625–635.

Greifeneder, E. (2016). The effects of distraction on task completion scores in a natural environment test setting. *Journal of the Association for Information Science and Technology, 67*(12), 2858–2870.

Grimes, C., Tang, D., & Russell, D. M. (2007). Query logs alone are not enough. *Workshop on query log analysis at WWW*.

He, J., & Yilmaz, E. (2017). User behaviour and task characteristics: A field study of daily information behaviour. *Proceedings of the 2017 ACM SIGIR conference on human information interaction and retrieval* (pp. 67–76). ACM.

Hendahewa, C., & Shah, C. (2017). Evaluating user search trails in exploratory search tasks. *Information Processing & Management, 53*(4), 905–922.

Jansen, B. J. (2006). Using temporal patterns of interactions to design effective automated searching assistance. *Communications of the ACM, 49*(4), 72–74.

Jansen, B. J., & McNeese, M. D. (2005). Evaluating the effectiveness of and patterns of interactions with automated searching assistance. *Journal of the American Society for Information Science and Technology, 56*(14), 1480–1503.

Jiang, J., He, D., & Allan, J. (2014). Searching, browsing, and clicking in a search session: Changes in user behavior by task and over time. *Proceedings of the 37th international ACM SIGIR conference on research & development in information retrieval* (pp. 607–616). ACM.

Kafai, Y., & Bates, M. J. (1997). Internet Web-searching instruction in the elementary classroom: Building a foundation for information literacy. *School Library Media Quarterly, 25*(2), 103–111.

Kelly, D. (2009). Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval, 3*(1—2), 1–224.
Kelly, D., Arguello, J., Edwards, A., & Wu, W. C. (2015). Development and evaluation of search tasks for IIR experiments using a cognitive complexity framework. *Proceedings of the 2015 international conference on the theory of information retrieval* (pp. 101–110). ACM.
Kelly, D., & Azzopardi, L. (2015). How many results per page? A study of SERP size, search behavior and user experience. *Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval* (pp. 183–192). ACM.
Kelly, D., & Belkin, N. J. (2004). Display time as implicit feedback: Understanding task effects. *Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 377–384). ACM.
Kelly, D., Cushing, A., Dostert, M., Niu, X., & Gyllstrom, K. (2010). Effects of popularity and quality on the usage of query suggestions during information search. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 45–54). ACM.
Kriewel, S., & Fuhr, N. (2007). Adaptive search suggestions for digital libraries. *International conference on Asian digital libraries* (pp. 220–229). Berlin, Heidelberg: Springer.
Li, Y., & Belkin, N. J. (2008). A faceted approach to conceptualizing tasks in information seeking. *Information Processing & Management, 44*(6), 1822–1837.
Liu, J. (2017). Toward a unified model of human information behavior: An equilibrium perspective. *Journal of Documentation, 73*(4), 666–688.
Lucchese, C., Orlando, S., Perego, R., Silvestri, F., & Tolomei, G. (2013). Discovering tasks from search engine query logs. *ACM Transactions on Information Systems (TOIS), 31*(3), 14.
Moraveji, N., Russell, D., Bien, J., & Mease, D. (2011). Measuring improvement in user search performance resulting from optimal search tips. *Proceedings of the 34th international ACM SIGIR conference on research and development in information retrieval* (pp. 355–364). ACM.
Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the Web. *Journal of the American Society for information Science and Technology, 53*(2), 145–161.
Schutz, A., & Luckmann, T. (1973). *The structures of the life world.* Evanston, IL: Northwestern University Press.
Scott, T. J., & O'Sullivan, M. K. (2005). Analyzing student search strategies: Making a case for integrating information literacy skills into the curriculum. *Teacher Librarian, 33*(1), 21–26.
Shah, C. (2018). Information fostering: Being proactive with information seeking and retrieval. *Proceedings of the 3rd ACM SIGIR conference on human information interaction and retrieval (CHIIR)* (pp. 62–71). ACM.
Suls, J., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science, 11*(5), 159–163.
Syed, R., & Collins-Thompson, K. (2017). Retrieval algorithms optimized for human learning. *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval* (pp. 555–564). ACM.
Tanford, S., & Montgomery, R. (2015). The effects of social influence and cognitive dissonance on travel purchase decisions. *Journal of Travel Research, 54*(5), 596–610.
Teevan, J., Alvarado, C., Ackerman, M. S., & Karger, D. R. (2004). The perfect search engine is not enough: A study of orienteering behavior in directed search. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 415–422). ACM.
Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology, 37*(4), 571–590.
Wang, Y., Liu, J., Mandal, S., & Shah, C. (2017). Search successes and failures in query segments and search tasks: A field study. *Proceedings of the Association for Information Science and Technology, 54*(1), 436–445.
Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology, 53*(2), 134–144.
Wilson, P. (1983). *Second-hand knowledge: An inquiry into cognitive authority.* Westport, CT: Greenwood Press.
Xie, I., & Joo, S. (2010). Transitions in search tactics during the Web-based search process. *Journal of the American Society for Information Science and Technology, 61*(11), 2188–2205.
Xu, J., & Croft, W. B. (1996). Query expansion using local and global document analysis. *Proceedings of the 19th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 4–11). ACM.
Yamamoto, Y., & Yamamoto, T. (2018). Query priming for promoting critical thinking in web search. *Proceedings of the 3rd ACM SIGIR conference on human information interaction & retrieval* (pp. 12–21). ACM.