Delft University of Technology

How Does Team Composition Affect Knowledge Gain of Users in Collaborative Web Search?

Xu, Luyan; Zhou, Xuan; Gadiraju, Ujwal

DOI
10.1145/3372923.3404784

Publication date
2020

Document Version
Accepted author manuscript

Published in
HT'20: Proceedings of the 31st ACM Conference on Hypertext and Social Media

Citation (APA)
Xu, L., Zhou, X., & Gadiraju, U. (2020). How Does Team Composition Affect Knowledge Gain of Users in Collaborative Web Search? In *HT'20: Proceedings of the 31st ACM Conference on Hypertext and Social Media* (pp. 91-100). Association for Computing Machinery (ACM). https://doi.org/10.1145/3372923.3404784

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

How Does Team Composition Affect Knowledge Gain of Users in Collaborative Web Search?

Luyan Xu
DEKE Lab(MOE),
Renmin University of China
Beijing, China
xuluyan@ruc.edu.cn

Xuan Zhou
School of Data Science & Engineering,
East China Normal University
Shanghai, China
xuan.zhou@outlook.com

Ujwal Gadiraju
Delft University of Technology
Delft, The Netherlands
u.k.gadiraju@tudelft.nl

ABSTRACT

Studies in searching as learning (SAL) have revealed that user knowledge gain not only manifests over a long-term learning period, but also occurs in single short-term web search sessions. Though prior works have shown that the knowledge gain of collaborators can be influenced by user demographics and searching strategies in long-term collaborative learning, little is known about the effect of these factors on user knowledge gain in short-term collaborative web search. In this paper, we present a study addressing the knowledge gain of user pairs in single collaborative web search sessions. Using crowdsourcing we recruited 454 unique users (227 random pairs), who then collaboratively worked on informational search tasks spanning 10 different topics and information needs. We investigated how users' demographics and traits, and the interaction between these factors could influence their knowledge gain. We found that in contrast to offline collaboration cases, user demographics such as gender, age, etc. do not significantly effect users' knowledge gain in collaborative web search sessions. Instead, our results highlight the presence of labor division of queries and particular interaction patterns in communication that facilitate knowledge gain in user pairs. Based on these findings, we propose a multiple linear regression model to predict the knowledge gain of users in collaborative web search sessions from the perspective of team composition.

CCS CONCEPTS

• **Information systems → Collaborative search**; • **Human-centered computing → User studies**; *User models.*

KEYWORDS

Collaborative Web Search, Knowledge Gain, Team Composition

ACM Reference Format:
Luyan Xu, Xuan Zhou, and Ujwal Gadiraju. 2020. How Does Team Composition Affect Knowledge Gain of Users in Collaborative Web Search?. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media (HT '20), July 13–15, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
HT '20, July 13–15, 2020, Virtual Event, USA
© 2020 Association for Computing Machinery.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Although web search has been largely considered as a solitary activity, recent studies of web search habits have revealed an increased prevalence and frequency of collaborative search, in particular among younger users [2, 26, 29]. On one hand, the context of collaborative web search has gained renewed interest in the field of search as learning (SAL), wherein remote users collaborate synchronously to search information and gain knowledge about particular topics [16, 26]; collaboration can also benefit users' searching process and search results in web search sessions [29]. On the other hand, it has been found that learners' collaborative learning tends to converge with the activity of web searching, that users search independently on the web and coordinate using separate communication tools [14, 26, 27]. Recent work has acknowledged the benefits in interacting with even strangers during collaborative searching process to obtain different perspectives [40]. It is at this confluence that the potential of collaborative web search has been recognized.

In situations where learning is facilitated through collaborative web search, it is fundamental to first understand how users' knowledge gain is effected through collaborative web search. An important step towards solving this puzzle pertains to how collaborators can be paired or teamed up, and how their intrinsic characteristics and searching strategies affect the learning related outcomes of individuals in the pair or team. The majority of collaborative web search is in group size of two people [25], which has been resonated by findings from later studies [22, 23]. There have been many efforts investigating the effect of strategies and user characteristics on knowledge gain of users during a long-term period such as a semester or several weeks [8, 20, 41]. However, little is known about the learning outcomes of users involved in short-term collaborative web search sessions. With advances in technology, short-term collaboration in web search sessions between remote users with the same information need is a realistic possibility. The effect of user demographics, traits and collaboration strategies on their knowledge gain are important, yet unanswered questions.

In this paper, we address this problem by investigating the strategies adopted by different searcher populations (e.g. elder / younger, expert / non-expert, college-educated / non-college educated etc.) to understand how users' knowledge gain can be affected by their characteristics and searching strategies. We focus on paired remote users having the same information need, and synchronous collaboration in web search. We make original contributions by addressing the following research questions:

RQ1: How do user demographics and individual traits affect their performance in collaborative web search sessions? We investigate whether the user demographics and individual traits that have been demonstrated to be influential in offline or long-term collaborative web learning have an effect on searchers' learning outcomes in single web search sessions. We found that in contrast to other collaboration settings, works exploring collaborative web search sessions with remote users working spontaneously on topics have undermined the affect of user's demographics (i.e. age, gender) in their knowledge gain.

RQ2: What strategies do successful user pairs adopt to maximize their knowledge gain in collaborative web search sessions? We investigated the searching strategies of collaborators and their effect on knowledge gain. We found all the participants in our study cooperated on tasks as peers performing balance search activity, rather than adopting roles of a leader and executor; the labor division was naturally elicited among all user pairs while the negotiation about it in the communication happened in most(87%) of user pairs. Pairs that exhibited high knowledge gain were more active and fluent in sharing discoveries and topical discussion, issued more diversified queries and located target webpages more efficiently.

RQ3: How can the searchers' intrinsic characteristics and the team strategies affect their quality of SAL outcomes? We built a multiple linear regression model to investigate how the user demographics and their strategies in collaboration can consequently result in knowledge gain of individuals and pairs. The model highlights the following variables as knowledge gain predictors: users' education level, their domain knowledge, the overlap in query terms of a user with the terms in partner's query, the number of quick-back clicks and the proportion of sharing factual and topical discussion in their communication.

2 RELATED LITERATURE

We position our contributions with respect to three main realms of closely related work – (i) collaborative search, (ii) teaming strategies in collaborative learning; (iii) evaluation of knowledge gain.

2.1 Collaborative Search

Collaborative search [16, 34] is a subset of social search in which participants work together to satisfy an information need. Existing works have found that the collaboration between searchers can provide a number of benefits for improving learning experience, such as enabling participants to achieve greater recall [35], offering the potential to improve search skills through exposure to others' behavior [24], and providing an opportunity to strengthen social connections [22]. Collaborative search has been studied in situations in which collaborators are searching synchronously versus asynchronously, and co-located versus remote [34]. In this paper, we focus on remote and synchronous collaboration in web search which is common in today's web search environment [25].

The most prominent approach has been to develop dedicated systems for collaborative search [3, 11, 21, 28, 32]. All of these dedicated systems include a search interface, as well as peripheral tools for collaborators to communicate, share information, and gain awareness of each other's activities. However, recent research on both user behavior in lab experiment [28] and recent real-world collaborative search practice [14, 26] reveal that while people often search in group, they do so without using the features from dedicated systems such as automatic division of labor. Instead, they search independently and coordinate using light-weighted communication tools such as instant messaging. Following this trend, we developed a collaborative search system called *PairSearch* consist of a searching platform and a separate communication panel to support our experimental study in this work.

2.2 Teaming Strategy in Collaborative Learning

Academic investigation in area of collaborative learning (CL) and computer-supported collaborative learning (CSCL) have studied the influence of team composition strategy on the searching and learning outcomes. Savicki et al [33] revealed that the composition of groups with different education levels and genders of students could be closely related to the ways their learning performance and outcomes. Webb et al. [39] conducted course-based experiments among students, focusing on analyzing how the factors of group members affects the group and individual learning outcomes. They found that students in high-ability (high education level) in teams tend to contribute more by providing more explanations; besides, low-ability students with high-ability peers as teammates are more likely to significantly improve their performance on both group tasks and individually-completed post-test. Recent works like [12] and [4] applied the features of group members like gender, knowledge background or skills, and education levels,etc. and conducted models for the evaluation of composition strategies and the prediction of knowledge gain through long-term collaboration.

2.3 User Search Behavior and Knowledge Gain

Some existing studies have focused on the correlation between learning progress and individual user activity features. Eickhoff et al. [6] investigated the distinct evolution of particular features throughout search sessions and the correlation of document features with the actual learning intent. The influence of distinct query types on knowledge gain was studied by Collins-Thompson et al. [5], finding that intrinsically diverse queries lead to increased knowledge gain. Gadiraju et al. [10] described the use of knowledge tests to calibrate the knowledge of users before and after their search sessions, quantifying their knowledge gain.

By matching the learning tasks into different learning stages of Anderson and Krathwohl's taxonomy [19], Jansen et al. studied the correlation between search behaviors of 72 participants and their learning stage [15]. They showed that information searching is a learning process with unique searching characteristics corresponding to particular learning levels. Gwizdka et al. [1] proposed to assess learning outcomes in search environments by correlating individual search behaviors with corresponding eye-tracking measures. Syed and Collins-Thompson [37] proposed to optimize the learning outcome of the vocabulary learning task by selecting a set of documents that consider the keyword density and domain knowledge of the learner.

We extend the current understanding of users' knowledge gain by focusing on collaborative informational search sessions. By simulating real world information needs and search sessions on the Web, we present an analysis of quantifiable knowledge gain in collaborative searching scenarios.

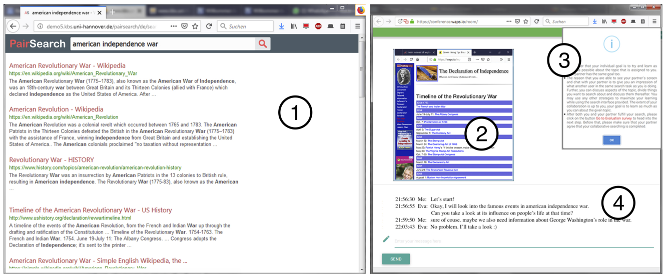


Figure 1: Collaborative searching using *PairSearch*

3 METHOD AND EXPERIMENTAL SETUP

We developed *PairSearch*, an online collaborative search platform, and recruited users from a crowdsourcing platform.

3.1 System Design and Rationale

Following the most recent prototype mentioned in related work, we built an online collaborative search platform called *PairSearch*, with an aim to support collaboration between two remote users searching the web synchronously with an intention to acquire knowledge pertaining to a given topic. As shown in the thumbnail (cf. Figure 1), the platform includes a browser window (Figure 1-①) and a separate light-weight collaboration panel (on the right). Within the browser window, we built a search system on top of the Bing Web Search API. We logged user activity on the platform including queries, clicks, keypresses, etc. using PHP / Javascript and the jQuery library. The collaboration panel has three major components: a real-time screenshare (Figure 1-②), a chat-box (Figure 1-④), and a pop-up window providing "Instructions" (Figure 1-③) when it is needed. This platform is available online [1].

Design Rationale for Process and Product Collaboration
Coordinated searching and sharing facts are important search tactics observed in previous work [38]. We integrated collaboration functions into a separate panel from the searching window since recent works revealed that users prefer separate lightweight tools for communication during collaborative search [14, 26]. Through the chatbox (Figure 1-④) users could not only communicate but also share resources (i.e. links, images, etc.) with their partners; by using the screenshare panel (Figure 1-②) users can see their partner's real-time searching activities. We note that by providing visual access to the collaborator's search activities, we overcome the frustrations regarding a lack of awareness of collaborators' activities and the resulting redundant work as showcased in previous work by Morris et al. [26]. We did not implement advanced features such as automatic labor division and split search since previous works have found that these advanced features were used rather rarely in the collaborative search process [28].

When two users are paired (i.e. when they are randomly allocated to the same team), they can search independently, directly see each other's search activities such as the executed queries, visited pages, mouse movements, etc. Note that only the search activities of users within the browser window (i.e. Figure 1-①) are shared by the screenshare function, and any private information is neither detected nor logged.

Table 1: Topics and corresponding information needs presented to user pairs in the collaborative informational search sessions.

Topic	Information Need
1. Altitude Sickness	In this task you are required to acquire knowledge about the symptoms, causes and prevention of altitude sickness. (20 items)
2. American Revolutionary War	In this task, you are required to acquire knowledge about the 'American Revolutionary War'. (10 items)
3. Carpenter Bees	In this task, you are required to acquire knowledge about the biological species 'carpenter bees'. How do they look? How do they live? (10 items)
4. Evolution	In this task, you are required to acquire knowledge about the theory of evolution. (12 items)
5. NASA Interplanetary Missions	In this task, you are required to acquire knowledge about the past, present, and possible future of interplanetary missions that are planned by the NASA. (20 items)
6. Orcas Island	In this task you are required to acquire knowledge about the Orcas Island. (20 items)
7. Sangre de Cristo Mountains	In this task, you are required to acquire knowledge about 'Sangre de Cristo' mountain range. (10 items)
8. Sun Tzu	In this task, you are required to acquire knowledge about the Chinese author Sun Tzu - about his life, his writings, and his influence to the present day. (15 items)
9. Tornado	In this task, you are required to acquire knowledge about the weather phenomenon that is called 'tornado'. (20 items)
10. USS Cole Bombing	In this task, you are required to acquire knowledge about the 2000 terrorist attack that came to be known as the 'USS Cole bombing'. (10 items)

3.2 Specifying Information Need

We employed the 10 topics and corresponding information needs (see Table 1) to study knowledge gain of users in informational search sessions that have been used and corroborated in previous work by Gadiraju et al. [10]. The topics were randomly selected from the *TREC 2014 Web Track* dataset[2]. Note that knowledge on all topics was measured using scientifically formulated knowledge tests created by Yu et al. [42] comprising between 10 and 20 items. The answer options were in all cases 'TRUE', 'FALSE', and 'I DON'T KNOW'. The differences in the number of items reflects the varying scopes of the information needs; relatively narrow (e.g., *Carpenter Bees*-10 items) as well as broad (e.g., *NASA Interplanetary Missions*-20 items).

3.3 Participants

We recruited 500 participants from Prolific[3], a premier crowdsourcing platform. To ensure that users can understand the instructions and interact with partners fluently, we recruited workers only from English-speaking countries or marked English as first language by the platform [9]. To undermine the platform effect [31] in our study,

[1] *Pairsearch*:https://conference.waps.io/?uid=ht_test

[2] http://www.trec.nist.gov/act_part/tracks/web/web2014.topics.txt

[3] Prolific: https://www.prolific.co/

Table 2: User demographics.

Gender	female / male	
Age	18-25 / 26-35 / 36-45 / 46-55	
Education	**under college**	**college and upper**
	No schooling Some high school, no diploma High School Some college, no degree Technical/trade/vocational training Associate degree	Bachelor's degree Master's degree Doctorate degree

we recruited users from Prolific who also have accounts and are active in other popular platforms such as MTurk or FigureEight. Related studies have found that the user demographics such as gender, age and education level could be influential factors in users' collaboration with their partners in offline or long-term learning cases. In this work, we mark participants by these demographics to represent diverse groups of online users as shown in Table 2. Participants were paid in accordance to minimum wage regulations, in addition to individual bonuses (in total, participants who participated in our study were paid £1, 450). We interchangeably refer to these participants in our experiment as users henceforth.

4 STUDY DESIGN

4.1 Experiment Procedure

At the onset, participants were informed that the task entailed '*searching the Web for some information collaboratively*'. In practice, the experimental sessions lasted for at most 40 minutes and were structured as described below (with the average duration):

1. Tutorial (2 min): Participants first read a short set of tips to help them learn about the platform and the basic functionalities required during the task.
2. Pre-test knowledge calibration (3 min): Participants were asked to respond to a few questions (technically referred to as '*items*'[4]) corresponding to a particular topic, without searching the Web for answers. Items took the form of statements pertaining to a topic, and participants had to select whether the statement was 'TRUE', 'FALSE', or 'I DON'T KNOW' in case they were uncertain. To encourage the participants to respond without external consultation, we informed them that their responses to these questions would not affect their pay. We also encouraged participants to provide responses to the best of their knowledge and avoid guessing. Results of this pre-test were used to calibrate the knowledge of individual participants with respect to the assigned topic.
3. Information need description (2 min): Participants were presented with an information need corresponding to the topic of the calibration test they completed (in accordance with previous work; cf. [10]).
4. Lounge (2 min): Participants were randomly paired with other participants who completed the calibration test pertaining to the same topic (i.e. with the same information need) in time sequence through a circular queue and allocated a unique room_id. Each pair of participants were asked to take a few minutes to communicate and get to know each other.
5. Collaboration (20 min): Pairs were told to use *Pair-Search* to collaboratively search the Web and learn as much as possible about the topic, satisfying the prescribed information need.
6. Post-task knowledge calibration (3 min): Each pair were informed that to successfully finish the task, they would individually have to complete a final test on the topic, when they felt that they were ready. Furthermore, participants were conveyed the message that depending on their accuracy on the final test they could earn a bonus payment. Our rationale behind this was to incentivize a genuine zeal to learn among the participants.

We subsequently logged all the activities of the participants within the *PairSearch* platform. Participants were encouraged to proceed to the next stage only once they felt that their information need was satisfied and when they were ready for the post-session test. On completing the post-session test, users received a unique code that they could enter on the Prolific platform to claim their reward.

4.2 Data Collection

To ensure the reliability of responses and the behavioral data that is produced in the search sessions, we filtered pairs of users using the following criteria:

- Users who entered no queries in the *PairSearch* system meanwhile did not communicate with their partners. Since the aim of our work is to investigate users' knowledge gain and behaviors in collaborative web search sessions, we discard those users who neither enter a search query nor communicate with partners.
- Users (or their collaborators) who selected the same option; either 'TRUE'/'FALSE' or 'I DON'T KNOW', for all items in the knowledge calibration test or the post-session test.
- Users (or their collaborators) who did not complete the post-session test.

We filtered out 46 users using this criteria, resulting in 454 users i.e. 227 unique pairs of users across the 10 topics. The analysis and results presented hereafter are based on these 454 users alone.

4.3 Measuring Knowledge Gain

We measure the knowledge gain of users in collaborative search sessions corresponding to a given information need as the difference between the pre-session knowledge calibration score and the post-session knowledge calibration score[5]. To avoid potential ceiling effects caused by the limited questions in the questionnaires, we normalize the knowledge gain of users by applying the ratio between absolute knowledge gain and the maximum possible value of knowledge gain that users can achieve, which is commonly used in the field of Education and Learning [13, 30]: $g_N = (postCalib - preCalib)/(1 - preCalib)$, wherein g_N represents the normalized knowledge gain, and $postCalib$ ($pre-$) refers to the post-session (pre-) knowledge calibration score (in %).

4.3.1 *Knowledge Gain in Collaborative Search vs. Single-user Search.*

Prior work has explored the knowledge gain of users in single-user search sessions using identical tasks and setup [10]. It was found that nearly 70% of the users ($N = 420$) exhibited a knowledge gain after searching. In this work, we found that around 85%

[4] Scientifically formulated knowledge tests corresponding to different topics and information need in Table 1 in comprise these items. The knowledge tests have been shown to have high internal reliability (using Cronbach's α) in [10].

[5] We consider 'I DON'T KNOW' options that were selected, as incorrect responses while computing the knowledge calibration scores and post-session test scores.

of individual users and 92% of user pairs exhibited a knowledge gain across all topics of tasks and search sessions. We note that users in this work generally exhibited a knowledge gain (normalized) of 44% ($M = 43.97, SD = 29.22$). In comparison to prior work [10], we found that collaboratively searching to satisfy information needs lead to a higher absolute knowledge gain ($postCalib = 67.45, preCalib = 32.35$) of around 35% than that observed in single-user search sessions ($postCalib = 57.77, preCalib = 38.22$) of around 19%. A two-tailed T-test between single-user search and collaborative search on the absolute knowledge gain across ten search tasks and all search sessions revealed that this difference is statistically significant ($t(873) = 3.92, p < .01$). These results confirm that collaborative search yields a greater increase in knowledge in comparison to single-user search.

4.3.2 Task Effect on Knowledge Gain. To understand the effect of topical differences in tasks on users' knowledge gain, we conducted a one-way between users ANOVA. We did not find any significant effect of topics on the knowledge gain of users. We thereby do not control for topical differences when analyzing the knowledge gain of users further in this paper.

5 USERS IN COLLABORATIVE WEB SEARCH

In this section, we investigate the role of user demographics and traits in gaining knowledge from collaborative web search sessions.

5.1 Demographics and Domain Knowledge

5.1.1 Demographics. Among all the users participated in the experiment, 67% were male. 45% of users were aged 18 - 25; 27% of users were aged 26 - 35; 20% of users were aged 36 - 45; 8% of users were aged 46 - 55. Of all the users, 40% receive college or higher education (i.e. 25% had a college degree, 14% had a master degree and 1% had a doctorate degree), while 60% did not (i.e. no participant reported they received no schooling; 3% had received some high school education yet no diploma; 16% had a high school degree; 5% received technical / vocational training; 33% had completed some college yet no degree; 4% had an associate degree). All users in our experiment reported they used search engines frequently and several times per day.

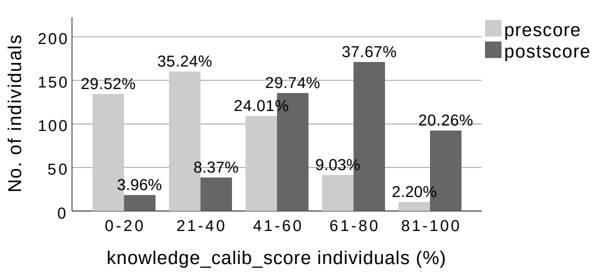


Figure 2: Distribution of knowledge calibration scores in pre- / post-session knowledge tests.

5.1.2 Domain Knowledge. We explore the impact of users' domain knowledge by leveraging their pre-session knowledge calibration scores. As shown in Figure 2, before collaborative searching the majority of users exhibited a pre-session knowledge test score of < 40%, indicating a moderate to low domain knowledge on the given information needs across the 10 topics.

We categorize users with regard to their level of domain knowledge (DK) by applying *Standard Deviation Classification* approach. Given the approximately normal distributions of knowledge calibration scores(X), we transformed (X)-scores into Z-scores with a mean of 0 and a Standard Deviation (SD) of 1. We used the statistically defined intervals ($X < -0.5SD$ = low; $-0.5SD < X < 0.5SD$ = moderate; $0.5SD < X$ = high) for the classification of users into groups with low(L), moderate(M) or high(H) level of domain knowledge as shown in Table 3.

Table 3: Groups of users created based on their calibrated domain knowledge (DK); $mean \pm 0.5SD$.

	Mean (%)	SD (%)	Low	Moderate	High
DK	33.03	24.74	150	202	102

In view of the substantial variety of different topics, we argue that such a tripartite categorization of domain knowledge is meaningful and thus can be generalized to other similar intentional learning activities. This procedure weighs all knowledge tests equally irrespective of the number of items.

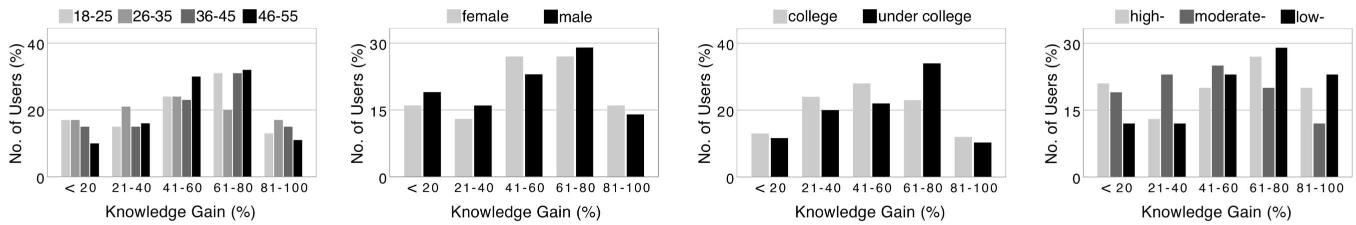
5.2 Knowledge Gain of Users

5.2.1 Knowledge Gain vs. User Demographics, Domain knowledge. We group users based on user demographics and levels of domain knowledge to analyze whether there's a group effect in the knowledge gain of users. Results are presented in Figure 3.

Figure 3a shows the overall distribution of users with different levels of knowledge gain across different age groups ranging from 18 to 55. We can observe that the majority of users exhibited a knowledge gain of 41% to 80%, while relatively older users (i.e. age of 46-55) learned more than the others. Using a one-way between users ANOVA, we found no statistically significant difference in knowledge gain of users across different age groups. Similarly, using Spearman's correlation coefficient we found no significant linear relationship between the knowledge gain of users and their age.

Figure 3b presents the distribution of users with respect to their gender. We did not find a significant effect of user gender on their knowledge gain. We note that these findings lie in contrast to those from previous studies revealing that younger [20] or older [18] people could perform better in gaining knowledge through long-term online collaboration, or female users [36] tend to be better in gaining knowledge from collaboration with others in online learning. These results suggest that although user demographics such as gender and age, have played a role in shaping users' knowledge acquisition, the effect is gradual and becomes significant over a long-term. On the other hand, there is no significant influence of user characteristics like gender and age in single web search sessions corresponding to collaborative learning.

Figure 3c presents the knowledge gain distribution of users who either had college or higher education and users who did not receive college education. We can see that the majority of both user groups exhibited a knowledge gain of 21% to 80%, while users who did not receive college education exhibited a relatively higher knowledge gain. Using a two-tailed T-test, we confirmed that this difference was statistically significant; the knowledge gain of college-edu



(a) Age (b) Gender (c) Education (d) Domain Knowledge

Figure 3: Distribution of knowledge gain of users belonging to different groups.

($M = 38.03, SD = 19.88$) is significantly lower than the not college-edu group ($M = 47.11, SD = 15.21$); $t(453) = 2.18, p < .05$.

As for the effect of domain knowledge (as shown in Figure 3d, we found that nearly half of the users who possessed a moderate domain knowledge exhibited a knowledge gain of 44% (the average knowledge gain of users in this experiment) or less, while certain proportion of users who possessed either a low or or high domain knowledge exhibited both very high(i.e. 81%-100%) or low knowledge gain (i.e. ≤ 20%). Using Spearman's correlation coefficient, we found a significant linear relationship between the knowledge gain of users and their domain knowledge; $R = -.62, p < .01$. An intuitive explanation for this observation is that the lesser a user knows about a topic, the more there is to learn through a collaborative search session, increasing the scope for knowledge to be gained through web searching.

5.2.2 Interaction between Education Level and Domain knowledge. Next we analyze how the education level and domain knowledge interactively effect knowledge gain of users in collaborative web search. To this end, we conducted a two-way ANOVA with the education (college education vs. under college education) and the domain knowledge (low- / moderate- / high- domain knowledge in section 5.1.2). We then applied Bonferroni corrections on the simple main effects test. Our results are presented in Table 4. We can see that there is a statistically significant interaction between education level and the domain knowledge on the knowledge gain of users. Simple main effect analysis showed the significant effect of domain knowledge on the knowledge gain of pairs with medium-large effect size ($\eta^2 \geq .05$). For users possessing low domain knowledge, we found that the knowledge gain of users who received college education ($95\%CI : (44.41, 73.68)$) is significantly higher than those who did not receive college education ($95\%CI : (32.67, 58.79)$). On the contrary, when possessing high or moderate domain knowledge, the knowledge gain of college-level users ($95\%CI : (13.06, 50.69)$) turned to be lesser than that of non-college-level users ($95\%CI : 29.44, 59.69$), though the difference is not statistically significant.

6 TEAM STRATEGIES TO GAIN KNOWLEDGE

By applying *Standard Deviation Classification* mentioned in section 5.1.2, we divide all pairs of users into 3 groups (low knowledge gain, 58 pairs with average knowledge gain of 5%; moderate knowledge gain, 92 pairs with average knowledge gain of 40%; high knowledge gain 77 pairs with average knowledge gain of 69%), and analyze the behavioral difference among these groups. We aim to understand the strategies that user pairs demonstrating a high knowledge gain adopt.

Table 4: Two-way ANOVA with education level and domain knowledge.

	F	p-value	η^2
Two-way ANOVA			
Knowledge Gain			
Education	21.133	p <.001	0.12
Domain Knowledge	14.373	p <.001	0.10
Education:Domain knowledge	8.81	p = 0.037	0.05
Simple main effects			
Domain Knowledge			
College-level	10.46	p <.001	0.05
Under College	12.349	p <.001	0.06
Education			
Low-Domain Knowledge	9.70	p <.001	0.08
Moderate-Domain Knowledge	3.00	p = 0.054	–
High-Domain Knowledge	1.83	p = 0.177	–

6.1 Searching Strategies

To understand how the three levels of knowledge gain of user pairs correspond to differences in their searching behavior, we investigate the nature of queries and clicks fired by the users across the groups.

6.1.1 Behavior Overview. We first report on the overall distribution of search behaviors in our experiment. We note that in each collaborative search session, users on average issued around 4 queries including more than 2 distinct queries, with the average query length of 4 terms. On average, users employed a minimum of 6 unique terms in their queries. We note consistency in the number of queries and clicks fired by collaborators within the same team – users in a pair issued similar numbers of queries and clicks, indicating a balance in their search activity. In each collaborative search session, a user on average fired 7 SERP clicks (around 2 clicks per query) and navigated to 5 distinct webpages other than SERP on average. The entire experimental session lasted for at most 40 minutes, during which users spent between 16 to 20 minutes on collaboratively searching the Web. On conducting a one-way ANOVA, we did not find significant differences in the basic characteristics of query (i.e. number. of queries; query interval; avg. query length) or the number of clicks across the three groups of user pairs.

6.1.2 Query Formulation and Search Splitting. Since users consumed the information prior to beginning the search session, we are interested in analyzing the fraction of query terms that go beyond the terms in the task description and knowledge tests. We

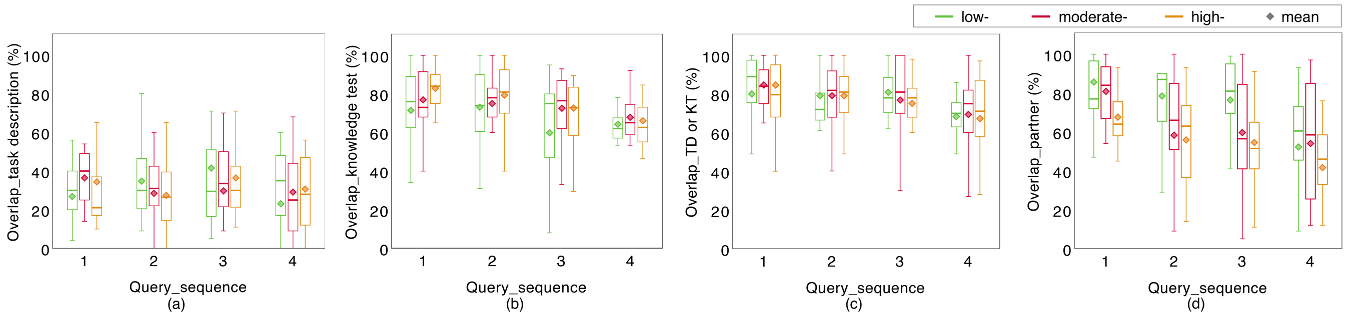


Figure 4: Overlap between the query terms issued by users with the terms in task description(TD), the terms in knowledge test(KT), the terms in TD or KT, and the terms in partner's queries across user quartiles based on knowledge gain.

analyze the overlap in query terms of pairs in total with the terms in task description, questions in the knowledge tests. Besides, as users work together on the same information need mainly through searching, we also look at how they split their search activity by analyzing the overlap in query terms of users with those of their partner's within the pair. We measure the overlap in query terms of a participant with the terms in task description, knowledge test and partner's query as the proportion of matched characters to the total length of the lemma of query terms (the query overlap of a user who issued no query during collaboration was counted as 1).

Figure 4 shows the evolution of the overlap between the query terms issued by users and the terms in task description, knowledge test and the partner's queries. As shown in Figure 4a, we found no significant difference among groups of pairs in the overlap between query terms issued by users and terms in the task description. While issuing queries, user pairs demonstrating a high knowledge gain tend to rely more on the information they consumed from knowledge tests than the other two groups at the first two queries fired on average (cf. Figure 4b, Query_sequence 1 and 2). A one-way between users ANOVA revealed that this difference is statistically significant ($p < .001$). In general while formulating queries, pairs in groups of moderate knowledge gain and high knowledge gain go beyond the information provided in knowledge test and task description as they progress in the task(cf. Figure 4c). This suggests that users employ new concepts or terms to search as they become more familiar with the information need and the collaboration. On the contrary, users in the low knowledge gain group did not demonstrate a discernible pattern.

As shown in Figure 4d, we note that all pairs of users exhibited a tendency to split the search activities as they progressed in the task. This labor division in queries is observed as users become more familiar with the information need and collaborative search. Compared to the other two groups of user pairs, high knowledge gain pairs were more adept at splitting their search activities in the first query itself (one-way ANOVA $p < .01$), suggesting that the high knowledge gain of users may partially arise from the efficient labor division in searching. Using Pearson's **R**, we found a negative linear relationship between the overlap of queries within collaborators and the knowledge gain of individuals' ($R = 0.42, R^2 = .18; p < .05$). This suggests that user pairs can benefit from splitting their search activities with their partners. Thus, labor division in collaborative search explains around 18% of the variance in knowledge gain of individual users' in our study.

6.1.3 Locating Information. To understand how the three groups of user pairs differ from each other in locating information, we analyze the SERP clicks fired by users in pairs. Prior works have shown that *quick-back clicks* (i.e. result clicks within a dwell time less than 10 seconds [17]) can be attributed to the difficulty experienced in locating target information. We thereby count the number of *quick-back clicks* on SERPs as a symbol of struggle in finding target information. In general, the number of *quick-back clicks* fired by the high knowledge gain pairs ($M = 2.25$) was less than the number of *quick-back clicks* fired by low knowledge gain pairs ($M = 3.75$), showing that the high knowledge gain pairs were better in locating target information to gain knowledge. By applying Kruskal-Wallis test, we found that this difference is statistically significant ($p = .02$).

6.2 Communication Strategies

In collaborative web search, communication is the main method for remote searchers to interact synchronously with each other [8]. In this study, users triggered a total of 7782 conversations across the different information needs. We note that users across different groups of knowledge gain entered 17 sentences on average to chat with their partners. We found no significant difference in the number or the length of conversational sentences issued by users across the three levels of knowledge gain. To understand whether and how communication between user pairs could effect their knowledge gain, we investigate the difference in communication across the user pair groups demonstrating a low, moderate and high knowledge gain respectively.

6.2.1 Interaction Patterns. We first analyze the pattern of communication between collaborators, aiming to understanding the effect of conversation mode on knowledge gain of users and knowledge transmission between user pairs. Prior work has revealed multiple types of interactions between collaborators [8]. For synchronous collaborative web searching, we consider the following three, focusing on the interaction between user pairs within each session:

- *Isolated interaction:* wherein a message posted by one user elicits no responses within a round of conversation;
- *Cyclic interaction:* wherein a message elicits only one response and a dialogue is set up between the initiator and the respondent within a round of conversation;
- *Chained interaction:* wherein a message elicits a series of messages within a round of conversation forming a chain.

Considering that different user pairs exhibit a varied preference for the time interval to reply to messages, we take the average response time of a pair as the threshold for determining a round of conversation and categorizing their interaction patterns. For example, given the average response time interval is 30 seconds, if a message sent by a user elicited no response in the average time interval, we then regard it as a round of *isolated interaction*, assuming that a new round of conversation would start after 30 seconds. Similarly, if a message sent by a user elicits several continuous messages from the partner within the time interval, we then consider it as a *chained interaction* between the user pair. Figure 5 presents the proportion

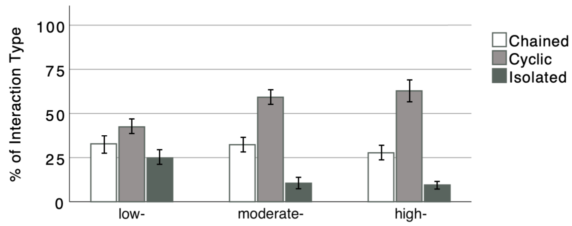


Figure 5: Proportion of the interaction pattern between collaborators in low- / moderate / high- knowledge gain pairs

of interaction patterns observed across the three groups of user pairs. We can see that *Cyclic interactions* are the most common interaction pattern of conversations across all pairs of users. Cyclic interactions occur significantly more often in the communication of high knowledge gain pairs and moderate gain pairs; one-way between pairs ANOVA, $p < .05$ level, $F(2, 226) = 4.27$.This suggests that fluid communication in collaborative search can lead to a higher user knowledge gain. Using Pearson's **R**, we found a weak positive linear relationship between average knowledge gain of pairs and the proportion of *cyclic interactions* in the corresponding communication; $\mathbf{R} = .29, R^2 = 08, p < .01$. Around 8% of the variance in the knowledge gain of users can be explained by the proportion of *cyclic interaction* in the users' communication within the search session. On the other hand, we found that *isolated interactions* occurs significantly more frequently (one-way ANOVA, $F(2, 224) = 7.88, p < .01$) between user pairs exhibiting a low knowledge gain than the other two groups. We did not find statistically significant differences in the occurrence of *Chained interactions* across the three groups.

6.2.2 Labor Division and Awareness Sharing. Next, we investigated how collaborators divide their labor and create an awareness of each other's activities through communication. To this end, we analyze the sentences in conversations between user pairs based on communication intent. Aligning with prior works [8, 28], we categorized the intent of collaborators' communication into 7 categories. Here, 1 and 2 relate to the negotiation of labor division and discussion of meta-issues, while 3 and 4 are primarily used in facilitating awareness [28]:

1. Planning search strategies and labor division (*e.g. "Let's divide the tasks..." "OK, I'll do the..." "Can you ..."*);
2. Discussing facts related to the general task topic (*e.g. "Altitude sickness could make me feel sleepy." "actually I learned this before, sun tzu is a strategist..."*);
3. Sharing facts discovered during the search (*e.g. "It says orcas island is the largest..., from wikipedia."*);
4. Sharing resources or links discovered during search (e.g."https://en.wikipedia.org /wiki/Orcas_Island");
5. Building relationships (*e.g. "Hello there...", "what's your name?"*);
6. Appreciation (*e.g. "Thanks a lot for..."*);
7. Others i.e. response or discussion not related to the topic (*e.g. "How can we check the instruction again?", "why", "okay", "agree", "will we get the bonus?" etc.*).

Table 5: Overall frequency of intent in communication.

Action	Number of pairs (and ratio)	Frequency	avg. proportion in conversations
labor division	197 (87%)	1332	0.17
meta-issues discussion	212 (93%)	1926	0.25
sharing information	201 (89%)	1188	0.15
sharing resources	135 (59%)	174	0.02
building relationships	227(100%)	624	0.08
appreciation	169 (74%)	437	0.06
others	227 (100%)	2101	0.27

We used crowdsourcing to classify the conversations. We dispatched sentences in the conversations as HITs to 100 workers from Mturk[6], providing them with adequate instructions and the seven categorization criteria. Table 5 presents the overall number and proportion of pairs in which users demonstrated a particular intent as well as the overall observed frequency of intents in our study. In the table, the 'avg. proportion in conversations' refers to the proportion of certain intents in all the conversations, we use this to show the general proportion that each intent could occupy in pairs' communication.

We observe spontaneous conversation between collaborators about their labor division for a given information need among most user pairs (87%). On investigating whether there is a relationship between the communication for labor division and the search split within user pairs, we did not find any significant relationship using Pearson' **R**; we also found no significant linear relationships between the proportion of conversations for labor division and the knowledge gain of users. These findings suggest that conversation about labor division itself does not directly determine the split in search activities or the knowledge gain of user pairs.

We found that nearly all user pairs (93%) discussed about their information need during collaborative search. User pairs demonstrating a high knowledge gain showed a significantly higher proportion of topical discussion than that observed in the other two groups (one-way ANOVA; $p < .05$ level; $F(2, 226) = 7.44$).

Sharing awareness (89%) was found to be even more prevalent among user pairs than labor division in our study. We note that most users preferred to share direct take-away facts that they discovered during search, rather than links or resources. Compared to the low knowledge gain user pairs, the moderate and high knowledge gain user pairs used significantly more proportion of their communication to share information they discovered during search (Mann-Whitney U, $p < .05$). From this we note that when communicating with each other, low knowledge gain pairs transmit less information they find during searching, which has an impact on their knowledge gain.

[6]https://www.mturk.com/

Table 6: Multiple Linear Regression with participants' education (Edu), Domain Knowledge (DK) and strategies in search and communication.

Category	Variables	β	Beta	95% CI	P value
	Individuals - Contant: 77.82				
Education	*Edu (UC as indicator)*				
	Edu (C)	-9.785	-0.113	(-12.993, -6.577)	<0.001
Domain Knowledge	*DK_In (L as indicator)*				
	DK_In (M)	-18.829	-0.296	(-23.793, -13.865)	<0.001
	DK_In (H)	-16.385	-0.202	(-20.172, -12.597)	<0.001
Search	*Overlap_KT (%)*	0.374	0.185	(0.183, 0.565)	<0.001
	Overlap_PQ (%)	-0.597	-0.437	(-0.815, -0.379)	<0.001
	No. of QB_clicks	-0.415	-0.282	(0.198, 0.632)	<0.001
Communication	*% of sharing fact*	0.210	0.113	(0.179, 0.241)	<0.001
	% of topical discussion	0.307	0.141	(0.159, 0.447)	0.0021
	Pairs - constant: 77.48				
Education	*Edu_Pair (UCUC as indicator)*				
	Edu_Pair (UCC)	-2.060	-0.177	(-3.977, -0.143)	<0.001
	Edu_Pair (CC)	-13.978	-0.197	(-16.283, -11.673)	<0.001
Domain Knowledge	*DK_Pair (LL as indicator)*				
	DK_Pair(ML)	-11.536	-0.420	(-17.364, -5.708)	0.019
	DK_Pair(HL)	-7.247	-0.381	(-10.681, -3.813)	<0.001
	DK_Pair(MM)	-13.079	-0.449	(-20.990, -5.168)	<0.001
	DK_Pair(HM)	-23.730	-0.228	(-31.943, -15.517)	<0.001
	DK_Pair(HH)	-26.446	-0.397	(-35.603, -17.288)	<0.001
Search	*Overlap_KT (%)*	0.342	0.241	(0.197, 0.487)	<0.001
	Overlap_PQ (%)	-0.654	-0.477	(-1.008, -0.299)	<0.001
	Avg No. of QB_clicks	-0.285	-0.273	(-0.421, -0.150)	<0.001
Communication	*% of sharing fact*	0.424	0.317	(0.281, 0.567)	<0.001
	% of topical discussion	0.326	0.213	(0.125, 0.527)	0.019

7 EFFECT OF PERSONAL CHARACTERISTICS AND TEAM STRATEGIES

From the previous sections, we have observed how different dimensions of user characteristics and team strategies can lead to varied levels of users' knowledge gain through collaborative search. In this section, we aim to understand the interactive effect of the these factors on collaborators' knowledge gain. Our findings can inform future research on how knowledge gain of individuals and user pairs in collaborative search sessions can be predicted or assessed from the perspective of team composition strategies.

We used Multiple Linear Regression analysis to develop models to predict knowledge gain of individual users and the average knowledge gain of user pairs. We regard all the variables that have been shown to have statistically significant correlations with knowledge gain in this work as candidate variables. For the categorical variable "Education" and "domain knowledge", we set "no college education"(*UC*) and "low domain knowledge"(*L*) as indicators and the other options as dummy variables in each feature. Similarly, we consider the notion for both users in a pair – "no college education"(*UCUC*) and "low domain knowledge"(*LL*) as indicators to predict the knowledge gain of pairs on average. Table 6 presents all the factors that are confirmed to be predictive on the knowledge gain of both individuals and pairs and with no multicollinearity, including: education level (*Edu*) – college education(C) or under college education(UC), domain knowledge (*DK*) – low(L) / moderate(M) / high domain knowledge(H); searching strategies – the overlap between query terms with terms in pre-session knowledge calibration test (*Overlap_KT*) and Partner's queries (*Overlap_PQ*), the number of *quick-back* clicks (*No. of QB_clicks*); communication strategies – the proportion of communication for sharing facts (*% of sharing fact, S*) and the proportion of communication for topical discussions (*% of topical discussions, D*) in communication. By checking the VIF and $Tolerance$ value, we note that there is no multicollineartiy among these features.

We found a significant linear relationship between the knowledge gain of individuals and their education, domain knowledge, search split as well the communication strategies($F = 25.019, p <$ .001), with an R^2 of .308, ($R^2{}_{adj} = .295$). Similarly, results also revealed a significant linear relationship between the average knowledge gain of user pairs and these features ($F = 27.548, p < .001$), with an R^2 of 0.292, ($R^2{}_{adj} = .271$). From this we note that nearly 30% of the variance in knowledge gain of individuals as well as that of user pairs in collaborative search sessions can be explained by the effect of team composition (i.e. the combination of collaborators education level and domain knowledge, and their searching and communication strategies). The knowledge gain of individuals and pairs (in %) can be represented by:

$$KG_i = 77.82 - \beta_{edu} * Edu - \beta_{DK} * DK + 0.374 * Overlap\_KT$$
$$-0.597 * Overlap\_PQ - 0.415 * QB\_clicks + 0.210 * S + 0.307 * D \quad (1)$$

$$KG_p = 77.48 - \beta_{edu} * Edu - \beta_{DK} * DK + 0.342 * Overlap\_KT$$
$$-0.654 * Overlap\_PQ - 0.285 * QB\_clicks + 0.424 * S + 0.326 * D \quad (2)$$

For example, we can interpret the features in the aforementioned equations as follows. For each percent increase in the facts shared from partners in communication, an individual's knowledge gain increases by a factor of 0.21. Users with a high-level of domain knowledge can obtain a knowledge gain score that is 16.385 less than users with low domain knowledge. With respect to the number of quick-back clicks in pairs, there is a decrease in the pair's average knowledge gain by a factor of 0.285 for each extra click.

To ensure that these features are generalizable to the whole population, and not just to the samples in this experiment, we adopted the cross validation method in [43]. For the knowledge gain of individuals (pairs), we divided all individuals (pairs) into two random subgroups and compared the estimated regression model for each subgroup with the final regression model in terms of each predictive feature. Results confirmed the explanation of the models to the knowledge gain of users.

8 DISCUSSION AND CONCLUSIONS

In this paper, we explored how user pairs in collaborative web search sessions gain knowledge and their use of different collaboration strategies. Our main findings show that: i In short-term collaborative web search sessions, the effect of user demographics (i.e. age, gender) on the knowledge gain is not significant, while the education level and domain knowledge of users can interactively effect users' knowledge gain to varying extents; ii Given a shared information need, labor division (i.e. search split) is naturally elicited between randomly paired remote online collaborators, and this behavior is conducive to users' knowledge gain; iii *cyclic interaction* is the most common interaction pattern of communication between users in collaborative web search, facilitates fluent conversations between users and aids in increasing knowledge gain; (iv) User pairs demonstrating a high knowledge gain benefit from a better performance in collaboration strategies such as splitting search activities, locating target information, smooth communication, and are more active in sharing awareness and topical discussion; (v) Users' education level, domain knowledge and their collaboration strategies have a significant linear relationship with their knowledge gain.

We make important contributions through this work. Our findings advance the current understanding of knowledge gain that manifests in short-term collaborative web search sessions. Moreover, by revealing the behavioral difference among user pairs across three levels of knowledge gain, our findings can inform future training and scaffolding strategies for helping users maximize their knowledge gain in collaborative web search. Through the multiple linear regression model, we found that user demographics, domain knowledge and the collaboration strategies between user pairs can collectively explain 30% of the variance in knowledge gain of both individuals and user pairs in collaboration. We note that this model fit is reasonably adequate given the complex nature of human characteristics and behaviors [7]. Following this type of research we can also envision advanced models based on user demographics and behavior data; for example, process mining that can be applied either to notify what will users do in the collaboration, or to predict knowledge gain of user / pair based on their characteristics and behavior. This would allow us to assess knowledge gain of users in collaborative web search after a few behavioral actions at the onset of collaboration, potentially resulting in broad implications on systems design for collaborative search. Finally, our work also demonstrates how crowdsourcing can be leveraged to study collaborative web search.

REFERENCES

[1] Anthony Aguirre, Matthew C Johnson, and Assaf Shomer. 2007. Towards observable signatures of other bubble universes. *Physical Review D* 76, 6 (2007).

[2] Sandeep Avula, Gordon Chadwick, Jaime Arguello, and Robert Capra. 2018. SearchBots: User Engagement with ChatBots During Collaborative Search. In *ACM CHIIR'18*. 52–61.

[3] Robert Capra, Annie T Chen, Katie Hawthorne, Jaime Arguello, Lee Shaw, and Gary Marchionini. 2012. Design and evaluation of a system to support collaborative search. *ASIS&T* 49, 1 (2012), 1–10.

[4] Ling Cen, Dymitr Ruta, Leigh Powell, Benjamin Hirsch, and Jason Ng. 2016. Quantitative approach to collaborative learning: performance prediction, individual assessment, and group composition. *International Journal of Computer-Supported Collaborative Learning* 11, 2 (2016), 187–225.

[5] Kevyn Collins-Thompson, Soo Young Rieh, Carl C Haynes, and Rohail Syed. 2016. Assessing learning outcomes in web search: A comparison of tasks and query strategies. In *ACM CHIIR'16*. 163–172.

[6] Carsten Eickhoff, Jaime Teevan, Ryen White, and Susan Dumais. 2014. Lessons from the journey: a query log analysis of within-session learning. In *ACM WSDM'14*. 223–232.

[7] Christopher J Ferguson. 2016. An effect size primer: a guide for clinicians and researchers. (2016).

[8] Yvonne YH Fung*. 2004. Collaborative online learning: Interaction patterns and limiting factors. *Open Learning: The Journal of Open, Distance and e-Learning* 19, 2 (2004), 135–149.

[9] Ujwal Gadiraju, Sebastian Möller, Martin Nöllenburg, Dietmar Saupe, Sebastian Egger-Lampl, Daniel Archambault, and Brian Fisher. 2017. Crowdsourcing versus the laboratory: towards human-centered experiments using the crowd. Springer.

[10] Ujwal Gadiraju, Ran Yu, Stefan Dietze, and Peter Holtz. 2018. Analyzing knowledge gain of users in informational search sessions on the web. In *ACM CHIIR'18*.

[11] Gene Golovchinsky, Abdigani Diriye, and Jeremy Pickens. 2011. Designing for collaboration in information seeking. *Proc. HCIR* (2011).

[12] Flavius L Gorgônio, Karliane MO Vale, Brazil Natal, Yuri KN Silva, and Huliane M Silva. 2017. Grouping Students for Cooperative and Collaborative Learning: challenges and trends in virtual learning environments. (2017).

[13] Richard R Hake. 1998. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American journal of Physics* 66, 1 (1998), 64–74.

[14] Marti A Hearst. 2014. What's missing from collaborative search? *Computer* 47, 3 (2014), 58–61.

[15] Bernard J Jansen, Danielle Booth, and Brian Smith. 2009. Using the taxonomy of cognitive learning to model online searching. *Information Processing & Management* 45, 6 (2009), 643–663.

[16] Ryan Kelly and Stephen J Payne. 2014. Collaborative web search in context: a study of tool use in everyday tasks. In *ACM CSCW'14*. 807–819.

[17] Youngho Kim, Ahmed Hassan, Ryen W White, and Imed Zitouni. 2014. Modeling dwell time to predict click-level satisfaction. In *ACM WSDM'14*. 193–202.

[18] Malcolm Shepherd Knowles. 1989. *The making of an adult educator: An autobiographical journey.* Jossey-Bass Inc Pub.

[19] David R Krathwohl and Lorin W Anderson. 2009. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.*

[20] Doo Hun Lim, Michael Lane Morris, and Seung-Won Yoon. 2006. Combined effect of instructional and learner variables on course outcomes within an online learning environment. *Journal of Interactive Online Learning* 5, 3 (2006), 255–269.

[21] Catalin Ioan Maican, Ana-Maria Cazan, Radu Constantin Lixandroiu, and Lavinia Dovleac. 2019. A study on academic staff personality and technology acceptance: The case of communication and collaboration applications. *Computers & Education* 128 (2019), 113–131.

[22] Winter Mason and Duncan J Watts. 2012. Collaborative learning in networks. *Proceedings of the National Academy of Sciences* 109, 3 (2012), 764–769.

[23] Felipe Moraes, Kilian Grashoff, and Claudia Hauff. 2019. On the impact of group size on collaborative search effectiveness. *Information Retrieval Journal* (2019), 1–23.

[24] Neema Moraveji, Meredith Morris, Daniel Morris, Mary Czerwinski, and Nathalie Henry Riche. 2011. ClassSearch: Facilitating the development of web search skills through social learning. In *ACM SIGCHI'11*. 1797–1806.

[25] Meredith Ringel Morris. 2008. A survey of collaborative web search practices. (2008).

[26] Meredith Ringel Morris. 2013. Collaborative search revisited. In *ACM CSCW'13*. 1181–1192.

[27] Meredith Ringel Morris and Saleema Amershi. 2016. Shared sensemaking: Enhancing the value of collaborative web search tools. (2016).

[28] Meredith Ringel Morris and Eric Horvitz. 2007. SearchTogether: an interface for collaborative web search. In *ACM UIST'07*. 3–12.

[29] Meredith Ringel Morris and Jaime Teevan. 2009. Collaborative web search: Who, what, where, when, and why. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 1, 1 (2009), 1–99.

[30] Jayson M Nissen, Robert M Talbot, Amreen Nasim Thompson, and Ben Van Dusen. 2018. Comparison of normalized gain and Cohen's d for analyzing gains on concept inventories. *Physical Review Physics Education Research* 14, 1 (2018).

[31] Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology* 70 (2017), 153–163.

[32] Sindunuraga Rikarno Putra, Felipe Moraes, and Claudia Hauff. 2018. Searchx: Empowering collaborative search research. In *ACM SIGIR'18*. 1265–1268.

[33] Victor Savicki, Merle Kelley, and Dawn Lingenfelter. 1996. Gender, group composition, and task type in small task groups using computer-mediated communication. *Computers in human behavior* 12, 4 (1996), 549–565.

[34] Chirag Shah. 2012. *Collaborative information seeking: The art and science of making the whole greater than the sum of all.* Vol. 34. Springer Science & Business Media.

[35] Chirag Shah and Roberto González-Ibáñez. 2011. Evaluating the synergic effect of collaboration in information seeking. In *ACM SIGIR'11*. 913–922.

[36] Glenda S Stump, Jonathan C Hilpert, Jenefer Husman, Wen-ting Chung, and Wonsik Kim. 2011. Collaborative learning in engineering students: Gender and achievement. *Journal of engineering education* 100, 3 (2011), 475–497.

[37] Rohail Syed and Kevyn Collins-Thompson. 2017. Retrieval algorithms optimized for human learning. In *ACM SIGIR'17*. 555–564.

[38] Michael B Twidale, David M Nichols, and Chris D Paice. 1997. Browsing is a collaborative process. *Information Processing & Management* 33, 6 (1997).

[39] Noreen M Webb. 1995. Group collaboration in assessment: Multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis* 17, 2 (1995), 239–261.

[40] Ryen W White. 2016. *Interactions with search systems.* Cambridge University Press.

[41] Amy Wojciechowski, Louann Bierlein Palmer, et al. 2005. Individual student characteristics: Can any be predictors of success in online classes. *Online journal of distance learning administration* 8, 2 (2005), 13.

[42] Ran Yu, Ujwal Gadiraju, Peter Holtz, Markus Rokicki, Philipp Kemkes, and Stefan Dietze. 2018. Predicting user knowledge gain in informational search sessions. In *The 41st ACM SIGIR*. ACM, 75–84.

[43] Xiangmin Zhang, Jingjing Liu, Michael Cole, and Nicholas Belkin. 2015. Predicting users' domain knowledge in information retrieval using multiple regression analysis of search behaviors. *ASIS&T* 66, 5 (2015), 980–1000.