A STUDY OF SOCIAL NAVIGATION SUPPORT UNDER DIFFERENT SITUATIONAL AND PERSONAL FACTORS

by

Rosta Farzan

B.S. in Computer Engineering, Sharif Univ. of Technology, 1999

M.S. in Computer Science, California State University, 2003

M.S. in Intelligent Systems Program, University of Pittsburgh, 2005

Submitted to the Graduate Faculty of

the Intelligent Systems Program in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

INTELLIGENT SYSTEMS PROGRAM

This dissertation was presented

by

Rosta Farzan

It was defended on

April 30, 2009

and approved by

Peter Brusilovksy, School of Information Sciences, University of Pittsburgh

Christian D. Schunn, Psychology, University of Pittsburgh

Daqing He, School of Information Sciences, University of Pittsburgh

Christine Neuwirth, Department of English, Carnegie Mellon University

Dissertation Director: Peter Brusilovksy, School of Information Sciences, University of Pittsburgh

A STUDY OF SOCIAL NAVIGATION SUPPORT UNDER DIFFERENT SITUATIONAL AND PERSONAL FACTORS

Rosta Farzan, PhD

University of Pittsburgh, 2009

"Social Navigation" for the Web has been created as a response to the problem of disorientation in information space. It helps by visualizing traces of behavior of other users and adding social affordance to the information space. Despite the popularity of social navigation ideas, very few studies of social navigation systems can be found in the research literature. In this dissertation, I designed and carried out an experiment to explore the effect of several factors on social navigation support (SNS). The purpose of the experiment was to identify situations under which social navigation is most useful and to investigate the effect of personal factors, e.g., interpersonal trust, and gender on the likelihood of following social navigation cues. To gain a deeper insight into the effect of SNS on users' information seeking behavior, traditional evaluation methodologies were supplemented with eye tracking. The results of the study show that social navigation cues affect subjects' search behavior; specifically, while under time pressure subjects were more likely to use SNS. SNS was successful in guiding them to relevant documents and allowed them to achieve higher search performance. Reading abilities and interpersonal trust had a reliable effect on the SNS-following behavior and on subjects' subjective opinion about SNS. The effect of the gender was less pronounced than expected, contrary to the evidence in the literature.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

To my parents who taught me to love learning

Acknowledgments

I would like to thank my advisor, Dr. Peter Brusilovsky, and the other members of my dissertation committee, Dr. Christian Schunn, Dr. Daqing He, and Dr. Christine Neuwirth for their valuable inputs, interesting ideas, and suggestions throughout my dissertation work.

I would like to thank Will Robb for proofreading my dissertation and Jaewook Ahn for the application support. The application used in this study is a branch of YourNews application developed by Jaewook Ahn.

This dissertation could not have been written without the support, encouragement, and friendship of Joan DiMicco, Behrang Mohit, and Tomek Loboda. I Particularly would like to thank Tomek Loboda for his help with statistical analysis, and all the valuable discussions we had.

1.0 INTRODUCTION

With information growing at an exponential pace, the information access tools that have served us well in the past are now creaking under the weight of the Web. Navigating through the ever-changing information space is becoming increasingly difficult, and even the latest search engine technologies are struggling to cope with our limited ability to declare our information needs. Recent research efforts have highlighted the interactive nature of information access behavior and promoted the potential value of harnessing user activity patterns to drive the next generation of information access tools. "Social Navigation" on the Web has been created as a response to the problem of disorientation in information space. Social navigation helps users by visualizing footprints of other users and adding social affordance about presence of others in information space [53], [7].

The idea of social navigation in information space stems from the natural tendency of humans to follow direct and indirect cues of each other when feeling lost. Different forms of social navigation have been demonstrated by animals or insects in nature, e.g. ants or birds foraging. Humans often follow social navigation in the physical world when dealing with making a decision, e.g. which movie to watch, which product to buy, which restaurant to go, etc. An interesting example of social navigation in the physical world is depicted in the picture in figure 1. Presence of a large number of people waiting for food, especially in cold weather, has served as an assuring clue about the quality of food at this place for the photographer [1]of this picture.

Although social navigation happens naturally in the physical world, careful consideration should be taken into account when transferring the idea into information space. It is essential to know how to collect and accumulate social information and how to present that informa-

[1]Thanks to Younghee Jung for sharing the photo (http://www.flickr.com/people/younghee/)

Figure 1: Example of social navigation in the physical world: “without knowing much, we joined the longest existing queue formed for a sushi restaurant. looking at faces of people (both young and old) filled with expectations despite the long wait in the cold weather, we were sure that the food would be worth every minute of waiting time. well, it was.”

tion to users to indeed guide them without overwhelming them with more information [7]. Additionally, it is important to know how and in what level users' navigation behaviors in information space is affected by social navigation. Are there any effects of external factors such as type and difficulty of the information seeking task or time constraint on usefulness of social navigation support? Is there a correlation between users' personal characteristics such as gender and adherence of social navigation cues?

In my earlier research [22], I studied the value of different types of social navigation support (SNS) in several applications in the educational domain. The effect of SNS on users' information seeking behaviors was evaluated in several classroom studies. The work attempted to present evidence that social visual cues affect students' browsing decisions and help them find relevant information faster. Classroom studies have the advantage of evaluating the system under the real situation and real usage of the system. However, very little can be controlled in classroom studies. Students are able to access the systems anytime and anywhere. Observing users' interactions with the system and their reaction to different features of the system is not possible. An access to an annotated link may be caused by the usefulness of the link, not the attached social cue. At the same time, the lack of an access does not really mean that the visual cue was not noticed - the link could be simply less relevant to the user in a specific context. The classroom studies provided broad perspective on the effect of SNS on students' information seeking behaviors.

In this dissertation, I designed and carried out an experiment to explore the role of several factors on the added value of SNS. The study is designed to identify situations under which social navigation is most useful, and investigate the effect of personal factors such as gender and interpersonal trust on following social navigation. To achieve deeper insight into the effect of SNS on users' navigation behaviors, I enriched traditional evaluation methodologies with eye tracking. A controlled experiment supplemented with eye tracking analysis allowed me to study some of the research questions important for design of social navigation in information space. Research questions such as: Do the users notice social navigation cues? Do the provided cues affect and change their link selection? Do the visual cues become more useful under time pressure when the user has little time to make a proper navigation decision? Is there an effect of gender, personal skills, and interpersonal trust on perception and adherence of SNS?

The main objectives of this dissertation are

- **To study the effect of social navigation support in information space in a controlled experiment**
- **To identify the effect of time constraint as an external factor on adherence of social navigation support**
- **To identify the effect of interpersonal trust and gender as personal factors on complying social navigation support**
- **To achieve deeper understanding of users actions by employing eye tracking**

1.1 SIGNIFICANCE

Despite the popularity of social navigation ideas, very few studies of social navigation systems can be found in the research literature [33]. The majority of research done in the field of social navigation can fall into the following categories:

- Conceptual structure: Focusing on theoretical discussion of social navigation phenomena and design aspects. This involves little or no focus on evaluation.
- Natural experiments: experiments that rely solely on observations of the effect of social navigation on the users' navigation behaviors in the system under study rather than manipulating variables in a controlled experiments

As a result, while there is a popular belief that SNS is powerful and helpful, little is known about the value of various social navigation approaches under different circumstances. This dissertation extends the previous work by studying the effect of social navigation support in a controlled experiment. Specifically I focused on the effect of time constraint, gender and interpersonal trust. In the following I present the motivation and significance behind studying the effect of these factors on following SNS.

1.1.1 Information Seeking and Time Constraint

As classified by [60], people apply three strategies to cope with time pressure: Acceleration, selection, and alteration of information search patterns. There is an indication of hierarchy of the strategies among people in which they start with acceleration, move to selection, and, if not enough, alternate their search strategy from alternative-based to attribute-based search patterns [64]. Weenig and Maarleveld [60] studied the effect of time pressure on complex choice tasks when people are dealing with large amounts of information. They argue that the effect of time constraint is particularly important for complex tasks. When dealing with complex tasks, people do not have the ability to examine all choices even when no time constraint is imposed, and time constraint magnifies this inability. They suggest that there is a curvi-linear relationship between attribute importance and attention level when performing complex choice tasks; i.e. the moderately important attributes are most significantly impacted by time constraint.

While the general effect of time constraint on information seeking behavior has been studied, the effect of navigation support and, specifically, social navigation support under time constraint has not been studied. The task studied in the current work can be considered a complex task, where users have to work with large amounts of information while considering different attributes of the information to complete the task. Social navigation support can help users cope with time constraint through strengthening selection strategy by providing extra information about more important information items. Social navigation cues are additional attributes about the information item.

I hypothesized that users who consider social navigation cues as crucial attributes pay more attention to those cues under time constraint.

1.1.2 Information Seeking and Interpersonal Trust

Trust plays an important role in Web Search and general information-seeking tasks. Golbeck [28] introduced the idea of Web trust and divided it into trust in content, trust in services, and trust in people. Trust in content refers to trust in information that the person sees. Credibility of the site, design and presentation of information, ease of finding information, technologies used to create the Web site, policy statement, and social recommendation are some of the factors influencing people's trust in content. Trust in services refers to the trustworthiness of Web services such as search engines and peer-to-peer systems. Trust in services are mainly judged based on performance of the service. Over time, people have built high trust in established Web search engines and they are highly likely to click on result ranked top in the list [36], [45]. Trust in people or interpersonal trust is "the evaluation of the trustworthiness of specific others such as managers and peers" [44]. Interpersonal trust can also refer to social trust in strangers.

It has been suggested that social navigation support can increase users' trust in the system by supplementing computer intelligence with human intelligence [15]. However, effectiveness of social navigation can be dependent on people's interpersonal trust levels. While the trust in services such as Web search has been studied, the effect of interpersonal trust on usefulness and compliance of navigation support has not been studied.

I hypothesized that users with high interpersonal trust are more likely to use social navigation cues as an attribute to evaluate the content while users with low interpersonal trust are more likely to ignore those cues.

1.1.3 Information Seeking and Gender

Navigation in hypermedia is closely related to navigation and wayfinding in physical and virtual world. Individuals with low spatial abilities have more difficulty with navigation in hypermedia [34]. Gender differences have been extensively studied in wayfinding literature. Males are known to exhibit better wayfinding performances; however, appropriate navigation support can improve female navigational performances and eliminate the difference [8]. It is also shown that different gender employ different strategies in wayfinding. Women are more likely to follow instructions on how to get from one place to another while men are more likely to use orientation of their own position in relation to environmental reference points [39]. Women are also more likely to make use of hand-drawn maps versus published road maps [39].

[19] provides a review of several studies of gender differences of information seeking

strategies in hypermedia. It has been found that females are more likely to feel disoriented and in need of navigation support while navigating in information space [25], [26]. Gender differences have been shown to play a role in Web search strategies as well. Females follow more of vertical search pattern, try longer queries, and spend more time evaluating returned Web pages and reading the text. On the other hand, males have more horizontal patterns, expect result with short (usually one word) queries, and quickly leave a returned Web page, but tend to explore more hyperlinks per minute [19].

Given the findings of gender differences in wayfinding in real world hypermedia navigation, I expected to observe gender differences in adherence of social navigation support. No prior work has studied whether different genders exhibit different patterns of following social navigation support.

I hypothesized that female participants are more likely to follow social navigation support.

1.1.4 Information Seeking and Eye Tracking

To supplement traditional log analysis and subjective evaluation, eye tracking is employed. Information in the log files are limited to click-streams and users' activities such as the order of viewing results, and their attention on social navigation cues cannot be captured in the logs. Eye tracking data provides information about users' areas of interest and attention and helps to closely examine the effect of social navigation cues on users' information seeking behaviors. More details on the usage of eye tracking for evaluation of information seeking tools is provided in chapter 2.

1.2 DISSERTATION ROADMAP

The remainder of this dissertation is organized as follows: In Chapter 2, I discuss the idea of social navigation and related work in the field. The chapter also covers a review of the research in eye tracking as a method for evaluating information access tools. Chapter 3 explains the research methodology and the design of the study. It explains the application used in the study, participants recruited for the study, and the experimental design, and data analysis consideration. Chapter 4 presents the result of the study in details and chapter 5 provides a discussion of the findings. Chapter 6 presents concluding remarks and limitations of the dissertation and discusses the impact of this dissertation on the field.

2.0 LITERATURE REVIEW

2.1 SOCIAL NAVIGATION

User navigation can be called social when it is driven by the actions from one or more advice providers [16]. It capitalizes on the natural tendency of people to follow direct and indirect cues about the activities of others. For example, people often prefer restaurants that appear to attract lots of customers or their movie preferences are often influenced by opinions of others. Over the last decade the idea of social navigation has attracted attention of researchers in information science as a way to offer navigational help in information spaces.

Social navigation assists users navigating through complex information spaces such as the Web by making use of latent community behavior. Social navigation in information space as well as the term social navigation was introduced by Dourish and Chalmers as "moving towards cluster of people" or "selecting objects because others have examined them" [16]. However, the idea of social navigation is frequently traced back to the pioneer Edit Wear and Read Wear systems [32]. Hill and Hollan introduce the idea of physical wear in the domain of document processing as"computational wear." Computational wear is the visualization of the history of authors' and readers' interactions with a document. The visualization of the history enables the new users to quickly locate the most viewed or edited parts of the document.

In summary, social navigation can help users in information spaces by: (1) Helping people to filter out relevant information and helping them to decide where to go next. (2) Adding social texture and quality to the information by presenting how other users react to the different information. (3)Adding social affordance to information space and making the space feel more inviting. (4) Providing more control to users and letting them reshape the information environment.

The literature exhibits different types of social navigation in terms of the communication mode, implementation approaches, and access methods. The communication modes define how the advice providers convey the advice to the users. Implementation approaches define how social navigation is designed and put into action. Social navigation was traditionally associated with browsing; however, it can enrich information seeking experience along different information access methods such as searching, and bookmarking. Figure 2 depicts a broad classification of social navigation in terms of communication modes, implementation approaches, and access methods.

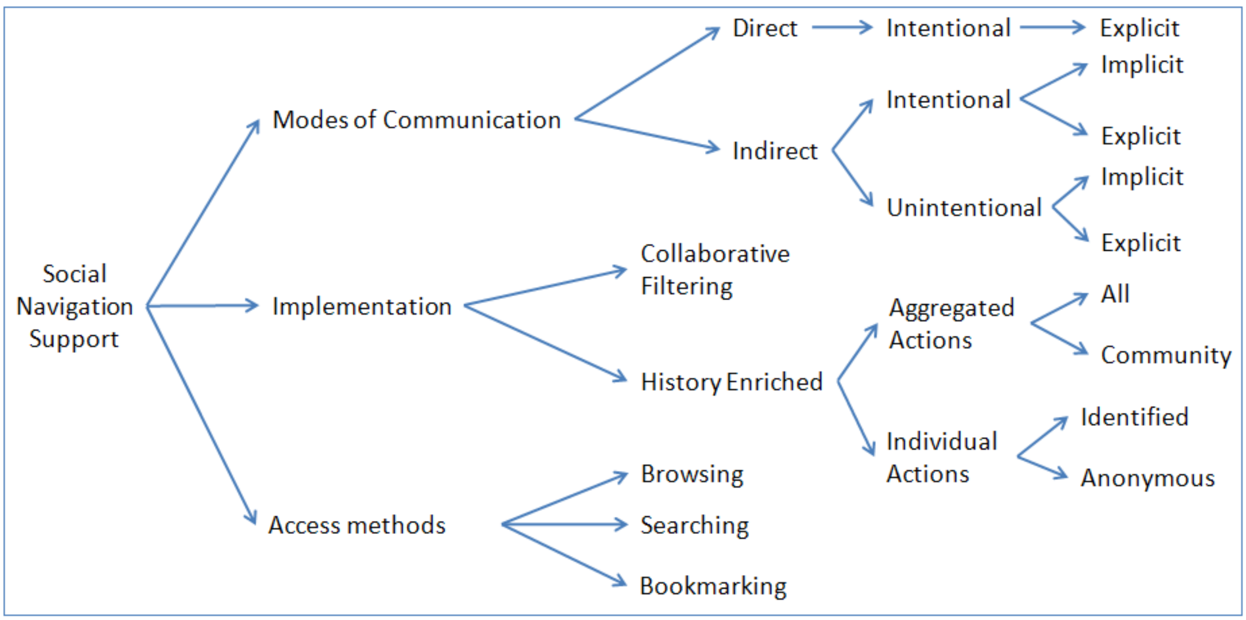


Figure 2: Taxonomy of Social Navigation in Information Space

2.1.1 Communication Modes

Dieberger [14] classifies social navigation into direct and indirect. He defines direct social navigation as synchronous and direct interaction between advice provider and receiver in the form of recommendation or guiding. Indirect social navigation focuses on aggregated history information such as wear traces in the Read Wear and Edit Wear systems. Furthermore

Svennson and Hook [59] suggest that communication of advice providers and navigators can be classified into intentional or unintentional. Direct social navigation happens in intentional way since the advice provider directly guides the navigator. However, indirect social navigation can happen in both intentional and unintentional way. An example of intentional and indirect social navigation can be a user who highlights part of a text to inform future users about what they had found important. The same action can be considered unintentional if the user highlights the text for personal matters and it is used by the designers of the system to guide other users. Additionally the feedback from advice providers can be collected in implicit or explicit mode. Keeping track of users' clickstream is an example of indirect and unintentional social navigation based on implicit feedback. Providing ratings for different objects such as movies can be considered example of indirect and intentional social navigation based on explicit feedback.

2.1.2 Implementation Approaches

Implementation of social navigation support (SNS) in information space can be classified into two main categories: collaborative filtering (CF) based recommender systems and history enriched systems. Collaborative recommender systems can be considered implementation of direct social navigation which change the navigation behavior of the users by promoting the relevant resources. Collaborative filtering is the process of refining information by considering preferences of users with similar interests. The rationale behind collaborative filtering is that users with similar goals and preferences makes similar decisions under analogous circumstances. CF based recommender systems have been implemented in several areas such as commercial products [1], movies [12], videos [31], or news [52]. The details about collaborative filtering and recommender systems are beyond this work. Refer to [54] for a comprehensive review.

History enriched systems are implementation of indirect social navigation. As suggested by Dieberger [14], they do not necessarily change users' navigation behavior but they increase their awareness. History enriched systems attempt to utilize traces of activities of early-on

[1] Amazon.com, Launch.com, CDNow.com

users to guide new users in the system. They visualize the aggregated or individual actions of a group of users [14]; for example, which of the available links were picked by the majority of similar users [61], or which pages are being explored by other users right now [38]. When dealing with individual actions the identity of users can be revealed or kept anonymous. While protecting user privacy is important and necessary, visibility and translucency can be beneficial to increase trust and awareness [18]. Aggregation of users' feedback can be limited to a specific community. This has the advantage of taking into account specific community interest while protecting the privacy of individual users in the community.

2.1.3 Information Access Methods

Social navigation is traditionally associated with browsing and the majority of early implementations of it were in the context of browsing. Social bookmarking and social search are more recent extensions of social navigation [42].

2.1.3.1 Social Browsing The systems Juggler [13] and Footprints [61] are classic examples that used social navigation to help users navigating in two kinds of information spaces - a Web site and a text-based virtual environment (MOO). Both systems attempted to visualize traces to guide future users. Footprints [61] provides history-rich navigation in complex information spaces such as the Web. Wexelblat and Mayes introduce the idea of interaction history for digital information which is taken from extensive human use of history traces in the physical world. They point out that not only a user's personal interaction history with a document, but also group history can be a useful navigation aid. Footprints provides contextualized navigation through the use of several interface features such as maps, path views, annotations, and sign posts. The system tracks all transitions from different sources such as selecting a link, typing a URL, or selecting a bookmark. It visualizes the interaction history by presenting the traffic through a Web site, percentage of users following each link, and popular paths to the Web sites. Additionally it allows the users to provide direct guidance by adding signposts expressing their opinions about different resources and the path to reach the resource. Juggler is an educational tool which combines a text-based virtual environment (known as MOO) and a Web browser. Juggler highlights major navigation paths through different textual bulletin boards (rooms) and adds the computational wear to each bulletin boards by showing the number of times it was accessed. Juggler also supports a direct form of social navigation by encouraging users to directly recommend useful resources (such as URLs) to each other.

Another example of a system with several forms of social navigation is KALAS [58], a food recipe system. It provides a history-enriched environment by visualizing the aggregated trail of users through the environment. The trail includes the comments left by the users as well as information about the number of users who have downloaded a recipe. Moreover, the system uses collaborative filtering to offer recommendation of recipes to users. To provide recommendation, KALAS collects users' feedback in an implicit and explicit format. Implicit feedback includes downloading, printing, or saving a recipe. Any of these actions will leave a positive vote for that recipe. Explicit feedback is collected by allowing users to click on a "good recipe" button or to check the thumbs-up/thumbs-down option in the recipe list. This provides an explicit positive or negative vote for the recipe. KALAS supports direct social navigation by displaying currently logged on users in each section of the system and allowing real-time chat among the users. KALAS has been evaluated by 302 users. The result of the evaluation shows that users make use of the recommendation feature very often and are very likely to be attracted to the most populated sections of the system; however, they were less influenced by the implicit trail left by other users and made little use of leaving comments.

2.1.3.2 Social Bookmarking Social bookmarking applications create a self-regulating collaborative network that lets users save, categorize, and share collection of bookmarks. The organization of resources inside these network is done through user-generated terms known as "Folksonomy." Users of social bookmarking applications can annotate links to resources with terms that are meaningful for them. These terms are known as tags. Tagging is the a bottom-up approach to classification which allows users to organize resources without any pre-determined classifications. It removes the barrier of hierarchically organized folder found in traditional browser based bookmarking by allowing the user to associate as many tags as they find relevant to a bookmark. Social bookmarking applications support reorientation of the focus of browsing from tags to users and vice-versa. This is known as “pivot browsing.”[43]. Along each direction, users are guided through SNS. They can easily find out how popular a resource is by the number of tags and the number of users who tagged and bookmarked a resource. Moreover, the popularity of tags can be determined by the number of times each tag is being used. “Tag clouds” are commonly used as a way to visualize high level tag popularity and provide tag based social navigation of resources. Information access using social tagging systems was recently popularized by such systems as Delicious [2], and Flickr [3].

CoFIND [17] is an example of an early social bookmarking system. It is a self-organized learning environment. All resources inside the system are entered by the learners. The classification and organization of the resources is done based on the feedback provided by the learners. They are able to explicitly rate the resources by associating different types of qualities (such as “simple,” “good for beginners”). The number of visits to each resource plays a role in the organization of the resources as an implicit form of feedback. Qualities are the same as tags in more recent social bookmarking systems such as Delicious.

Dogear [43] is a social bookmarking service designed for a large enterprise. Dogear is similar to Delicous but is adopted for large enterprise usage taking into account several design principles. The application requires real name identification to increase the social capital inside the enterprise by promoting resource and people discovery. The system supports both public and private bookmarks. It provides lists of recent and popular tag collections. Users can browse the public collections by tags, resources, or authors of the bookmarks. They can also add comments to the bookmarks. Millen et al [43] present how a large enterprise can also benefit from a social bookmarking service. Moreover, they discuss how the social bookmarking system can be integrated with other corporate collaborative applications.

2.1.3.3 Social Search Social search approaches attempt to improve user experiences with search engines by employing social information. Instead of focusing on search engine behavior, social search applications try to supplement search queries or search results with

[2]http://delicious.com/
[3]http://flickr.com

collective intelligence. They utilize past successful searches of similar users, search histories, and social annotations to guide users. Social search benefits from the presence of context in the from of queries. They can associate each result selection to a specific information need through the queries. As a result, interest of a different community of users can be defined by the type of their information need. Social search application helps users by pre-search or post-search guidance. Pre-search guidance is often done through query extension or suggestion. For example, if a user who is known to be interested in computer programming searches for "Java," they can be guided to expand their query to "Java programming" versus "Java coffee". Post-search guidance is frequently supported by annotating or re-ranking the search result [56], [1]. Search results can be augmented by the number of times a resource has been selected for a similar query, other relevant queries, or whether the page includes comments by similar users.

AntWorld is one of the pioneer social search systems that introduced the idea of integrating human intelligence as part of search technologies to increase the speed and accuracy of finding information on the Web [37]. AntWorld follows the ant metaphor and supports asynchronous collaboration among users. AntWorld defines information "quests" as series of queries pursued for a specific information need. Users are encouraged to express the information quests in natural language. The system allows the user to "judge" the relevance of pages they visit in relation to each quest. In addition to user explicit rating of relevance, the system keeps track of users' clickstream as an implicit indicator of interest. The relevant judgment is stored on the system centralized databases and used to recommend pages to future users with similar information quests. Recommendation is done by adding ant icons to links found relevant by other people within the quest.

I-SPY [56] offers another approach in improving search technologies with social navigation. I-SPY is a community based social search system that combines the re-ranking and annotating of search results to reflect the community interest while presenting the search results. The main idea behind I-SPY is that overlap of query terms represents similar information need. For each community of like-minded users, the system builds a hit matrix of query terms and successful search results for each query. I-SPY uses the hit matrix to promote resources that have been previously selected by other community users with similar queries. In addition to placing the most relevant result for the community at top of the result list, I-SPY annotates the result with visual icons representing popularity, related queries, and recency. The popularity icon shows the relevance of the result for the current query relative to the community in which it was entered. Its relevance value derives from the percentage of times this result has been selected by community members using the query. The related query icon informs the users of any related queries for this result; that is, other queries that have also led to the selection of this result by the current community. The recency icon provides users with information relating to the last time the result was encountered by users. This allows users to form a view of the freshness of the interaction trail. Results that have not been accessed in a number of months may not be as useful as results that have been accessed more recently. I-SPY has been evaluated through several user studies and presented promising results in improving users' search experiences.

2.1.4 Related Background Works

In my earlier works, I explored social navigation support in the context of educational information access in several applications. The applications focused on implementation of social navigation support based on different types of implicit and explicit user feedback while browsing or searching through the information spaces. The applications offered SNS taking into account different browsing behavior of past users such as number of visits, time spent reading, and users' annotations.

2.1.4.1 KnowledgeSea II KnowledgeSea II [5] is an educational information access tool. Knowledge Sea II is desinged to help students find relevant information among hundreds of online tutorial pages distributed over the Web. It provides SNS based on prior students' interactions with the system. It offers two types of SNS: Traffic-based and annotation-based. Traffic-based SNS is based on the classical footprints approach by counting how many users are passing through a link or visiting a page. The goal of traffic-based SNS is to recommend the most popular links and pages. However, the classic footprints approach can easily be challenged by snowball effect. KnowledgeSea II has taken into account the time spent reading a page to partially address the “snowball effect” problem [20]. Moreover, KnowledgeSea II employs users’ annotations as a more reliable indicator of user interest. Annotation-based SNS makes use of students’ annotations of educational resources. Students are encouraged to annotate Web pages they are reading by writing notes or highlighting parts of the page they found important. These annotations are stronger indicators of user interest and the importance of the page [21].

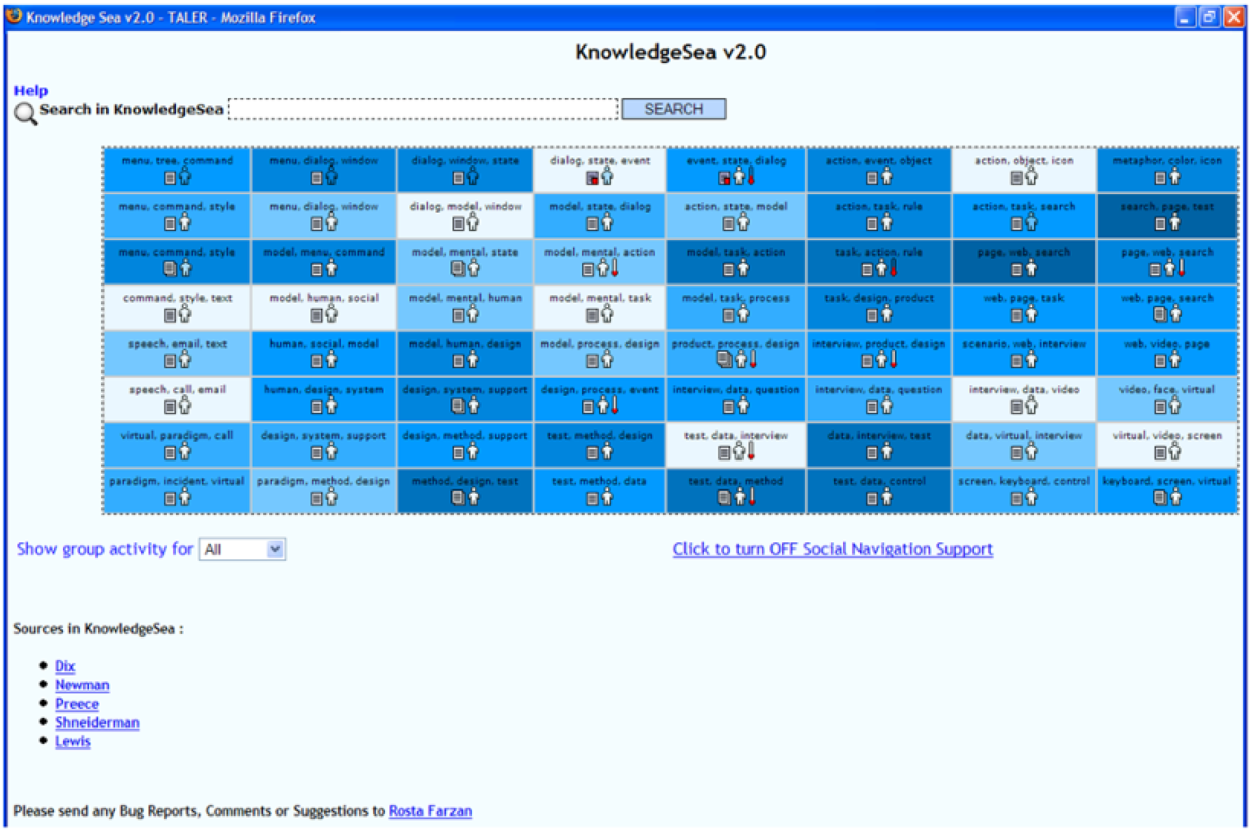


Figure 3: KnowledeSea II - map

The KnowledgeSea II system consists of three parts: the map, cell content, and resource page. The map is an eight by eight table of resources as shown in figure 3. Clicking on each cell will open a new window which shows the list of resources in the cell. Clicking on each item on the list will open the resource page (figure 4). The left side of the resource page includes the annotation frame which allows the students to associate a note with the page or highlight a specific part of the page. Annotations can be public or private, signed with students' names or anonymous, praise or general.

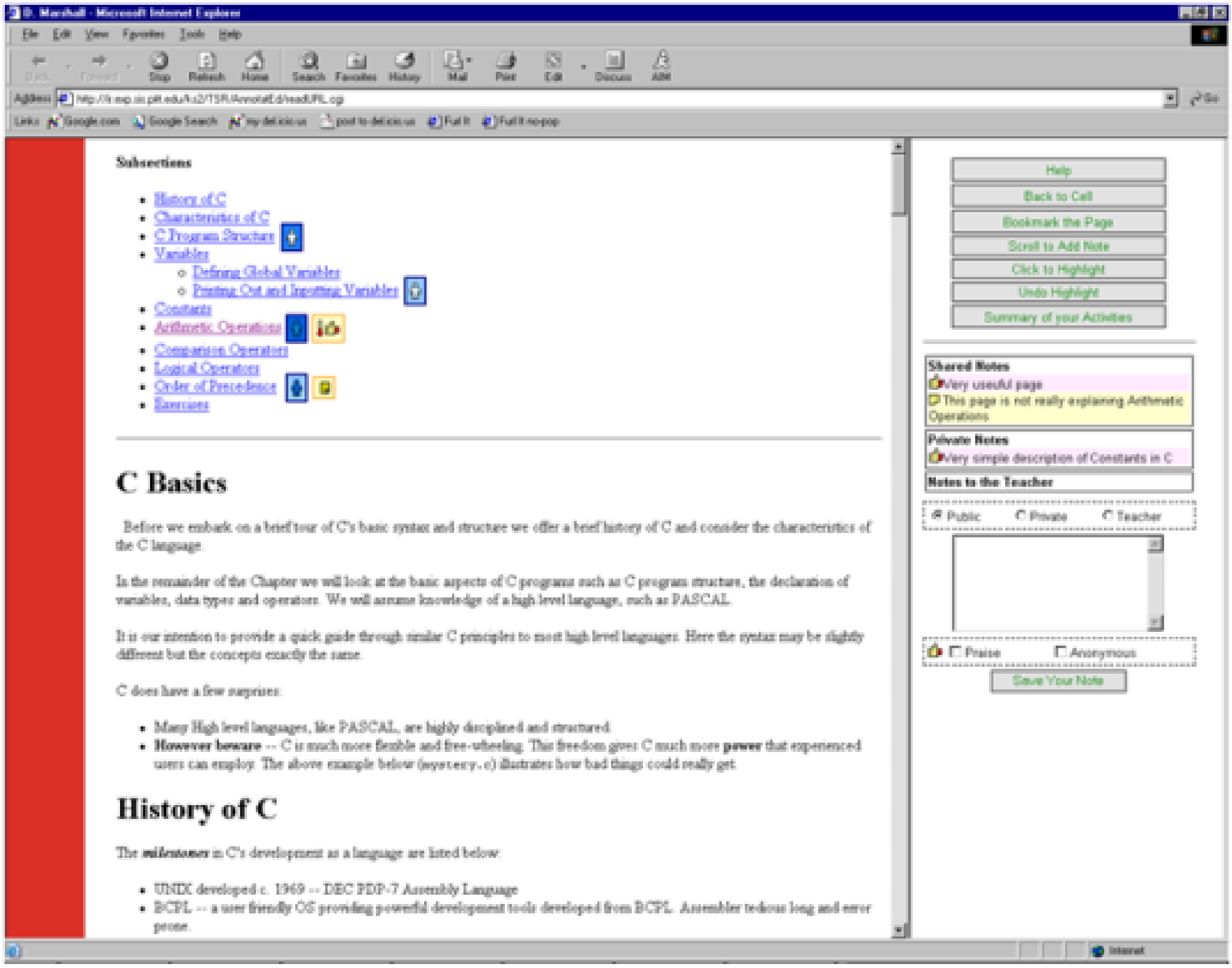


Figure 4: KnowledeSea II - Resource content page

Social navigation support is provided on all three parts of the system. The intensity of a cells background color on the map indicates how many users of the current group have visited each tutorial page and each cluster of tutorial pages. The more visits, the more intensive the color (Figure 3). On the cell content and resource page, all the links are annotated with an icon representing the same information. This kind of SNS helps the learners to clearly recognize the most and the least visited pages. Annotation based social navigation support is represented by visual icons [21]. The icon represents the type of the annotations and the magnitude of the annotations. On the map, cells including any resource with students' annotations will be augmented with a sticky note. Moreover, a thermometer on the cell represent the overall type of students' notes on the resources inside the cell. Warmer temperature represent more notes with a positive type. On the cell content and resource page, the links to resources with student's note are also augmented with the annotation icon.

The KnowledgeSea II system has been evaluated through several classroom studies. The result of the evaluation shows that users visit pages with social navigation cues more; specially the ones with annotation based cues. Also the result shows that stronger annotation based navigation support increases overall usage of the system and annotation activities [21].

2.1.4.2 Conference Navigator Academic conferences with multiple parallel sessions and high number of papers are an example of information overload. The conference navigator [23] system explores the value of social navigation in assisting the conference attendees to make decisions about which talks to attend. Conference Navigator addresses two critical issues related to the problem of collecting feedback from the users. First, feedback is achieved at a cost of users' time and attention. It often interferes with the natural order of their activities. Second, users are concerned about their privacy when they provide feedback. To deal with the first issue, the Conference Navigator attempts to introduce actions which provide reliable indication of users' interest while also being beneficial for them. The second issue is addressed by tracking activities at the community level instead of individual users. The collective wisdom is built based on feedback collected from a community of users with similar interests. The Conference Navigation system was evaluated at the E-Learn 2007 conference which is a good example of a large conference with several parallel sessions and large number of papers. The results indicate that the number of visits to the augmented papers with social navigation cues was higher than that of the non-augmented papers. The evaluation shows that users are likely to follow the footprints of the community and take advantage of social navigation support.

2.1.4.3 ASSIST-ACM The majority of existing social systems employ a single social technology such as social navigation, social bookmarking, or social search. Each system collects and exploits their own pool of community wisdom for the benefit of their users. ASSIST-ACM [24], [27] attempts to exploit the pools of wisdom from multiple social technologies, specifically social search and social navigation. The system was designed to provide community-based access to the articles in the Communications of ACM magazine. The search component uses the search history and browsing information to re-rank the returned result list so that it reflects the interests of the community. The browsing component uses the browsing and search history data to identify the pages that have been interesting to previous users. A classroom study of the system evaluated the general value of social browsing and social search and the added benefits of merging several sources of social wisdom. The result shows that social support helps the students to find more relevant articles in shorter amounts of time. Students were attracted to social cues and augmented links were more likely to be selected. Moreover, social search cues were highly utilized during browsing.

Figure 5 presents the classification of the aforementioned social navigation applications into my proposed taxonomy of social navigation.

2.1.5 Supporting Theories

Social navigation is inspired from principles that have been discovered in nature. People have observed a variety of interesting behaviors among insects or animals in nature. Animals and insects such as birds, fish, ants, or termites engage in collective or swarm behavior [40]. A swarm is a collection of non-sophisticated agents that are cooperating to achieve some goal. Each agent follows simple local rules from their environment in a relatively independent manner but collectively they achieve the swarm's objectives. This emergent collective intelligence is known as "Swarm Intelligence (SI) [3]." "SI is the property of a system whereby the collective behaviors of (unsophisticated) agents interacting locally with their environment cause coherent functional global patterns to emerge" [3]. An example of SI in nature is the food foraging behavior of ants. Ants use their pheromone to mark trails connecting the nest to food sources. The pheromone gets richer and richer as more ants follow the trail to carry food to the nest. At each point the trail with the highest density of the pheromone has the highest chance of being chosen by the ants.

	Communication					Implementation				Access Methods		
	Direct	Indirect				History Enriched						
		Intentional		Unintentional		Individual		Aggregated				
		Imp.	Exp.	Imp.	Exp.	Anon.	Iden.	Comm.	All	Browse	Search	Bookmark
Edit Wear & Read Wear				✓					✓	✓		
Footprints	✓			✓					✓	✓		
Juggler	✓			✓			✓		✓	✓		
KALAS	✓		✓	✓	✓		✓		✓	✓		
AntWorld			✓	✓					✓		✓	
I-SPY				✓				✓			✓	
CoFIND			✓	✓	✓	✓	✓	✓				✓
dogear			✓	✓	✓		✓		✓			✓
Knowledge Sea II		✓		✓		✓	✓	✓		✓	✓	
ASSIST-ACM		✓		✓		✓	✓	✓		✓	✓	
Conference Navigator		✓		✓		✓	✓	✓		✓	✓	
Dissertation Work				✓		✓		✓		✓	✓	

Figure 5: Classification of Social Navigation Systems

While interacting with complex information spaces, humans behave similar to animals in trying to achieve collective intelligence. Information seeking tasks on the Web can be mapped to a biological society. The Web represent the society, and the surfer represent the animal which is an autonomous agent with limited knowledge given the available information abundance. Desired information is food for which the surfer is browsing. Click-stream and other browsing behavior is the Web pheromone and the popularity of the Web page represents the density of the pheromone. Wu and Aberer [63] conducted a "Quest for Treasure" experiment to evaluate the collective intelligence behavior of humans in information space. The experiment involved 12 rooms that visitors could navigate to. Two of the rooms had a chest treasure in them. For each link they presented the raw visit click and pheromone density. Pheromone density was calculated taking into account positive and negative feedback. Positive feedback includes accumulation of visits and spreading of pheromone from other links. Negative feedback includes diffusion of the popularity of a link and was modeled by a half-life time function. Following the link pheromone, one could quickly find the treasure chests. The result of their experiment showed a simple form of self-organization and demonstrated the value of "swarm of internet surfers."

Effect of social navigation in information space can be explained by the information foraging theory. Related to SI, the information foraging theory [48] is an analogy to food foraging strategies among animals which states that "when feasible, natural information systems evolve toward stable states that maximizes gains of valuable information per unit cost." Information foraging is the result of human adaptation to the explosive information growth. The central problem the theory tries to address is allocation of attention to the most useful information. The goal is to maximize profitability of information resources by increasing information gained per unit cost. Information scent is used to assess the profitability of information resources. Information scent is the "perception of the value, cost, or access path of information sources obtained from proximal cues, such as bibliographic citations, WWW links, or icons representing the sources."

Information foraging has mainly focused on explaining information seeking behavior of individual users. Recently, Pirolli introduced the idea of "Social Information Foraging (SIF)." [47] SIF is based on the idea that information foragers engage in social exchange of information. Connected to the idea of swarm intelligence, information foragers cooperate to increase the likelihood of high-value information discoveries. The basic SIF model assumes existence of hints from the group of information forager about the likely location of useful information patches. It attempts to model the benefit of cooperation and social capital in information seeking tasks. Recent social Web technologies such as Web logs, collaborative tagging, and recommender systems have emerged to exploit or enhance SIF. The success of those technologies implies the effectiveness of social information foraging.

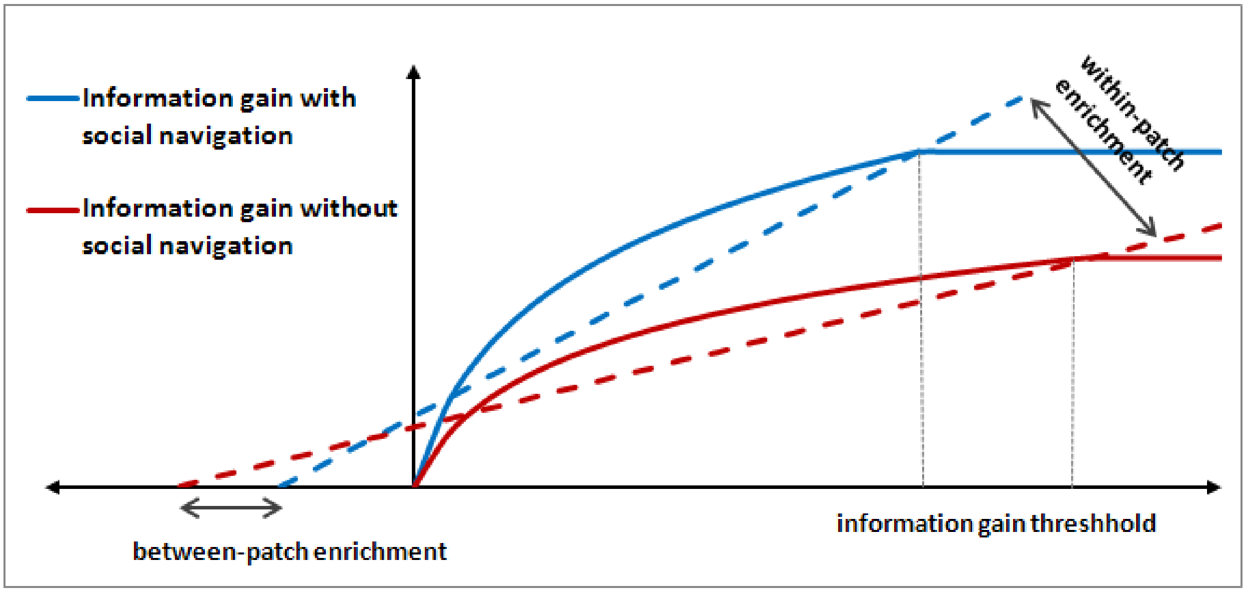


Figure 6: Information foraging model with SNS enrichment

SIF connects social navigation and information foraging. SNS can enrich information scent and assist in scent detection to judge the potential relevance of information resources. Information foragers have to navigate through information patches to find what they need. SNS can decrease the cost of information gain by both enriching between-patch and within-patch foraging gains. Figure 6 depicts the possible effect of SNS on information gain. To satisfy information needs, first, information foragers should find the relevant patches. As they go through the information patches they gain information as represented by the information gain function up to the point that they reach the information gain threshold. Social navigation cues can enrich between-patch information gain by highlighting the patches with useful information and decreasing the time needed to assess different patches. While navigating inside a patch, SNS can improve the return from a patch by highlighting the useful resources inside the patch; e.g. highlighting the part of the document that received the most attention by previous users.

2.1.6 Challenges of Social Navigation Support

Social navigation support is a successful and powerful approach in guiding users through complex information spaces such as the Web; however, it faces several challenges. Snowball effect, drift of interest, bootstrapping, and user participation are the most known challenges of implementing social navigation support in information space.

Snowball effect Social navigation relies on recommending the path traveled by others; however, if the first user heads in the wrong direction, all other users of the system can be attracted to the same wrong path. Therefore, it is important to be able to detect these paths and to prevent the system from directing users on to them. The snowball effect especially harms systems that rely mostly on implicit feedback from users. Combining several types of implicit feedback can partially address this problem; for example, combining time spent reading with clickstream data. If a user has gone through a page by mistake, the chance that they spend only a very short amount of time on the page is high. As a result, considering the time can help to eliminate some of the misleading pitfalls.

Drift of Interest A known challenge in implementation of social navigation is the concept of drift of interest [57]. Over time, the interest of people and the importance of information are changing. What is very important to a community of users today might not have much value in several months. This is especially important for highly dynamic context such as educational context in which the interest of students is dependent on the specific topic they are studying at the moment.

This problem can be addressed by weighing more recent visits, providing social navigation support based on the data from a specific period of time, or showing the recency of social guidance [56]. Often it is important to preserve old data in addition to recent ones. For example, in educational contexts, students might be interested in the currently discussed information to work on the latest assignment, and, at the same time, they might be interested in previously discussed materials to prepare for the midterm exam.

Bootstrapping A very important and well identified challenge in developing social navigation systems is how to get the system started. This is known as the "cold start" problem in collaborative filtering based recommender systems. Social navigation heavily relies on feedback provided by users - implicitly or explicitly. Early users will not have many navigational aids which might get disappointed by the system. On the other hand, as a result of not having navigational aid, they might head in the wrong direction which will affect the whole functionality of the system by accumulating a trail on the wrong path. Combining content based navigation support approaches with social navigation is the most common way of addressing the cold start problem [55].

2.2 EYE TRACKING AS A TOOL FOR ANALYSIS OF USER INFORMATION SEEKING BEHAVIOR

Traditional evaluation methods such as log analysis and usability testing provide limited information for understanding users' thought processes and their strategies while using a computer interface. An approach to enriching understanding of users' behaviors is employing eye movement data. Accurate viewing is only possible in 1-2 degrees of visual angle. As a result, gaze direction is a reliable indicator of the focus of attention. Rayner [50] has shown that eyes are attracted to the most informative areas of a scene because they are physically distinctive and informative. Eye movement data is typically divided into fixations and saccades. Fixations are relative pauses of eye movements over an informative region of interest for about 200-300 ms, while saccades are the rapid eye movements between the fixations with velocities as high as 500 degrees per second. Because of the quick eye movements, during saccades, no information is obtained [50].

The main methodology used in interface evaluation using eye-tracking is through dividing the interface into predefined areas of interest [29] and collecting users' eye movements on those areas. The number of fixations, the location of fixations, fixation duration, and cumulative fixation time are some of the most commonly used measures in the evaluation of computer interfaces using eye-tracking [49], [35]. Eyes usually fixate on areas that are interesting or important for the viewer. The number of fixations generally represent the amount of interest in an area with larger numbers for greater amounts of interest. Longer fixations represent more information processing that can be due to a higher density of information or more difficulty. Some additional measures have been used specifically to study users' information seeking behaviors. Scanpath length, transition matrix, and number of saccades are examples of those measures. Scanpath is the sum of the distance between gaze points. The scanpath is short if the information is well organized and the user can find the information easily. Transition matrix is the number of transitions between areas of interests. Higher numbers of transitions between areas represents difficulty of finding information. Higher numbers of saccades also represents difficulty of finding the desired information. More detailed discussion of eye movement metrics and what they measure can be found in [49] and [35].

2.2.1 Usage of Eye Movement in Evaluation of Information Retrieval Systems

More recently, researchers in the field of information retrieval have been attracted to using eye tracking techniques in Web search. Several works have utilized eye tracking for better understanding of users' search behavior and to model users and their interests beyond log analysis and queries they type in. The usage of eye tracking in information retrieval can be classified in terms of factors they study, measurements used, and methodology. Figure 7 presents a taxonomy of usage of eye tracking in Web search and information seeking studies.

Joachims et al [36] extended the work of assessing the reliability of implicit feedback by detailed analysis of users' decision making process through the use of eye tracking. They conducted an eye tracking study to evaluate the reliability of clickthrough data as implicit feedback. They analyzed users' fixations on a search result page to understand how to associate users' decision process with their clickthrough actions and how to generate feedback from clicks. Particularly, the eye tracking study allowed them to investigate the order users look at search results before they make a decision to click on a result. They showed that

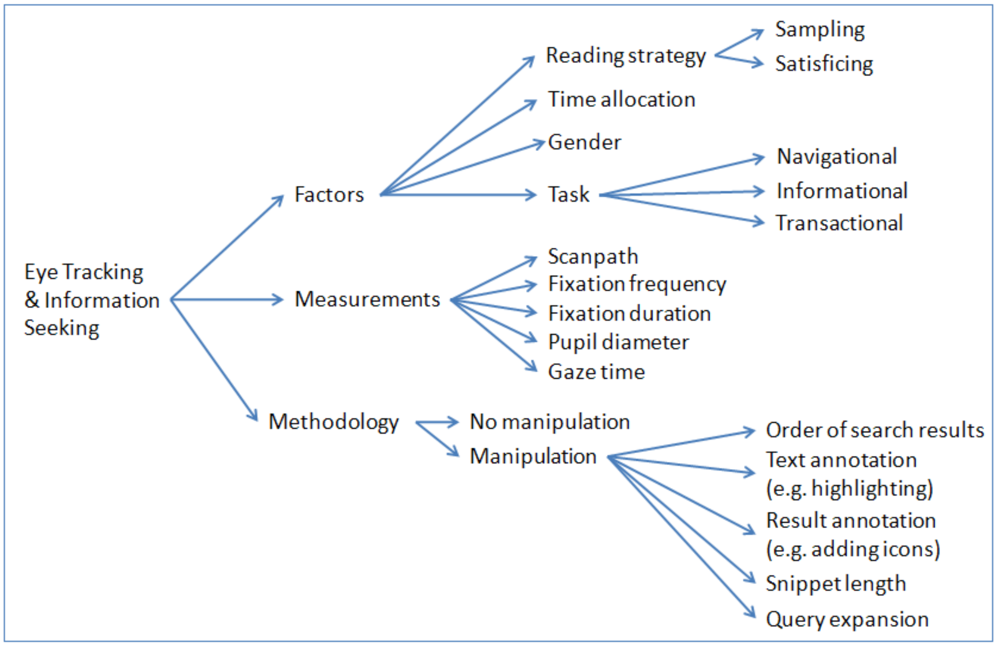


Figure 7: Classification of eye tracking studies in information retrieval systems

while first and second ranked results attract almost similar amount of attention, users are substantially more likely to click on the first result. The analysis of order of fixations also suggest that users tend to read the results in a linear order from top to bottom. In addition to studying the users' fixations pattern on usual search results, they manipulated the order of results to assess users' trust and quality biases. They define "trust bias" as bias toward clicking more on high ranked results, and "quality bias" as bias caused by the overall quality of the abstracts of search results. Their study confirms that users have substantial bias to click on links with high ranks and the quality of the ranking influences the users' clicking behaviors.

Broder [4] classifies search tasks into three different classes - navigational, informational, and transactional. Navigational task refers to the need to find a particular Web site. The intent behind informational tasks is to find specific information which can be found on one or more Web pages. The purpose of transactional queries is reaching Web sites for carrying out further transactions such as shopping or downloading various types of files. Lorigo et al [41] used eye tracking to study the effect of search task and gender on users' information seeking behaviors. They compared users' search patterns while performing navigational versus information search tasks. They conducted scanpath analysis and defined several scanpath properties to reason about search patterns. They observed that users rely heavily on the abstract to evaluate the relevance of a search result item. Very often the users evaluate the abstracts and choose to refine their search queries before they select any results. Their findings suggest that users generally do not follow the ranking presented by the search engine; they do few queries and examine the first few ranked abstracts. If the first two results do not seem to be relevant, then the users make jumps, skips, and regressions to check the rest of the abstracts. They observed significant effect of task on evaluation of the documents in terms of average time spent, average number of fixations, and average pupil dilation. Informational tasks take longer to be completed and users spent most of their time on documents and not on the results. While they did not observe a significant effect of gender on evaluation of the documents, there was a significant effect of gender on evaluation of search results. They observed that females make more regressions while males tend to be more linear in processing of the search results.

In a similar study, Cutrell and Guan [11] used eye tracking to study the usage of Web search. They conducted an experiment to study how users attend to different parts of Web search results and whether users' search strategies are different for navigational versus informational tasks. Specifically, they were interested in assessing the effect of snippet length on how people use Web search. They defined each individual search result as an area of interest and considered fixations longer than 100 ms in those areas of interests. Their result confirms the findings of the previous works that users mostly look at the first few search result items and they scan them in a linear order. Additionally, they observed a significant interaction of task type and snippet length. Longer snippets improved the success rate of informational search tasks while harming the navigational tasks. Eye movement data reveals that in both cases longer snippets draw more attention with the cost of less attention on the URLs. As a result, the performance of navigational search tasks that are focused on the URLs are being harmed.

Wilkinson and Payne [62] used eye tracking to study users' strategies in allocating time across multiple online texts under time pressure. Readers are known to employ two common strategies while reading: Sampling (quickly inspecting the text before making decision to read the text) or satisficing (continue reading until the quality drops below a satisfaction threshold) [51]. Wilkinson and Payne used eye tracking to achieve more insight on how readers employ satisficing strategy. They conducted experiments to ask participants to study for two tests and go over a set of texts under time pressure. Since the participants did not have enough time to carefully examine all the texts they had to make a decision to skip some texts. The analysis of participants' gaze data and scanpaths shows that while all participants followed satisficing strategy and did not reject reading a piece of text quickly, they did not read all of the text either. They read the beginning of a paragraph or text carefully and skimmed the rest to judge the relevance of the text.

Chi et al [9] studied the eye-gaze behavior of subjects to understand how highlighting keywords and sentences containing highly relevant conceptual keywords (ScentHighlights) affects subjects' reading behavior. They were interested in assessing whether highlighting is successful in directing users' attention while skimming the text. They analyzed users' initial fixations and eye behavior, and percentage of fixations on highlighted areas. Their results show that the users' initial fixations happen on highlighted areas and they fixate heavily on those areas. Their eye tracking evidence confirmed that users are more likely to pay attention to highlighted text.

Buscher et al [6] utilized gaze data in modeling users' short-term interests while working with a collection of documents. They take advantage of granularity of eye movement data to assess users' interest in specific parts of the documents. The result of evaluation of their approach shows that users' passage level attention as an implicit source of feedback can improve context-based query expansion and re-ranking.

2.2.2 Benefits of Adding Eye Movement Data to Log Data

In summary, eye movement data enhances the traditional log analysis evaluation approach in information retrieval in the following ways:

1. Scanpath analysis provides better understanding of the order searchers evaluate results.

2. Scanpath analysis and fixations data provide more accurate interpretation of implicit feedback such as clickthrough data.
3. Eye movement data provides detailed information about how users attend to different parts of search results such as a title or an abstract.
4. Eye movement data provides detailed information about what is viewed and what is skipped before making a click decision.

Figure 8 presents how current approaches in studying of information seeking using eye movement data fits into the proposed classification. The usage of eye tracking in this dissertation is similar to prior works in the field of information retrieval to understand users' information seeking strategies. The goal of this work is to employ eye tracking as a tool to uncover search strategies and how they vary under different conditions defined by presence of SNS and enforcement of time pressure. The bottom part of figure 8 shows the focus of the study in the current work. More details are discussed in chapter 3.

			Joachims et al	Lorigo et al	Cutrell & Guan	Wilkinson & Payne	Chi et al	Buscher et al	Dissertation Work
Factors	Reading Strategy					✓		✓	
	Time allocation					✓			✓
	Gender			✓					✓
	Task	Nav.		✓	✓				
		Info.	✓	✓	✓		✓	✓	✓
		Trans.							
Measurements	Scanpath		✓	✓	✓	✓	✓		
	Fixation duration								✓
	Fixation frequency			✓	✓		✓	✓	✓
	Pupil diameter			✓					
	Gaze time					✓	✓		
Methodology	No manipulation			✓					
	Manipulation	Reordering	✓					✓	
		Annotation					✓		✓
		Snippet len			✓				
		Query expansion						✓	

Figure 8: Classification of approaches in studying of information seeking using eye movement data

3.0 METHODOLOGY

To achieve the objectives discussed earlier I have conducted a user study explained in this chapter. This chapter covers the explanation of the task, the application used to perform the task, and the experimental design of the study.

3.1 TASK

The study investigated the effect of SNS in an "informational" task as defined by Broder's Web search taxonomy [4]. The participants were asked to respond to several questions by finding facts in a very large collection of relevant and irrelevant news articles from multiple sources. To encourage exploring and full usage of time, participants were particularly asked to find as much relevant information as time allows them. For each search task subjects were given a one-page task description providing a brief background to the task scenario and a list of questions to answer. A sample task description is shown in figure 9. For each topic, depending on the condition, participants were given seven or 15 minutes to search and collect notes that provided answers to the questions in the task scenario. At the end of the session, subjects were asked to associate notes they collected and questions. The question number(s) that a note provided response to was added to each note. A note could provide response to more than one question.

Topic – Earthquake hit India's Gujarat State

A huge earthquake hit India's Gujarat state, January 26, 2001. The task is to find the information about what rescue and relief actions have be taken on January 27, 2001, the second day after the earthquake.

From the articles, find snippets of relevant text to the following questions:

1. **Where did the earthquake happen?**
2. **What was the number of injure?**
3. **How many troops were sent in?**
4. **How many tents and other materials were needed?**
5. **How many relief materials were sent in?**

And here is the list of questions you colleagues have worked on:

1. What was the degree of the earthquake?
2. When did the earthquake happen?
3. What was the number of death?
4. How much money lost?
5. How about the lack of medical facilities?

Figure 9: Sample topic description for the information seeking task in the study

3.1.1 Document Set

The document collection used in this experiment was an expanded TDT4 (Topic Detection and Tracking)[1] test corpus. The corpus contains news articles from multiple sources in Arabic, Chinese, and English, and it covers events that happened between October 2000 to January 2001. Non-English articles are machine translated to English. The collection contains 28,390 English documents. Each TDT topic corresponds to a seminal event which happened at a specific time and place along with all necessary preconditions and unavoidable consequences. 18 of the original TDT4 topics are enriched into so-called GALE topics to resemble the tasks performed by information analysts. Each GALE topic contains an overarching task theme and up to 10 different but related factual questions [30].

[1] http://projects.ldc.upenn.edu/TDT4/

3.1.2 Ground Truth

As part of a bigger project, ground truth has been collected for 18 GALE topics. Ground truth for each GALE topic includes all the relevant documents in the collection. Relevant passages in each document is annotated by two human annotators under no time pressure. Each highlighted passage was marked as "slightly relevant" or "very relevant" by the annotators and was later associated with all the relevant questions in the topic by different annotators. Details about how ground truth was collected can be found in [30].

3.2 APPLICATION

I used a simple application similar to classical search engines for the study[2]. The main interface of the tool is presented in figure 10.

To use the application, users enter a query and the system returns a list of results ordered by relevance to the query. Each result page included 10 items and each item consists of a title and a small snippet of text as shown in figure 10. In addition to this traditional interface, the experimental version of the system offers two kinds of social navigation support. First, the search results are annotated with social navigation cues. The cues are based on two types of user activities: reading and highlighting. The human icon represents the amount of reading activity for the associated document and the annotation icon represents the amount of highlighting done in the document. The level of the filling color represents the density of the activity with a higher level of filling representing a higher number of activities. Mousing over the icons shows the details about the number of visits or number of highlights. Second kind of social navigation support is offered by "social maps." Social maps are two tables at top of the page, representing highlighting and visiting activities for the current 100 documents. Each cell in the table is associated with the document with the same rank as the cell number; i.e. the first one is associated with the first document in the list and so on. Users can directly

[2]This tool was originally developed by Jaewook Ahn [2]. I modified it for the purpose of this dissertation to accommodate social navigation support. Thanks to Jaewook Ahn for providing the access to the application and the code.

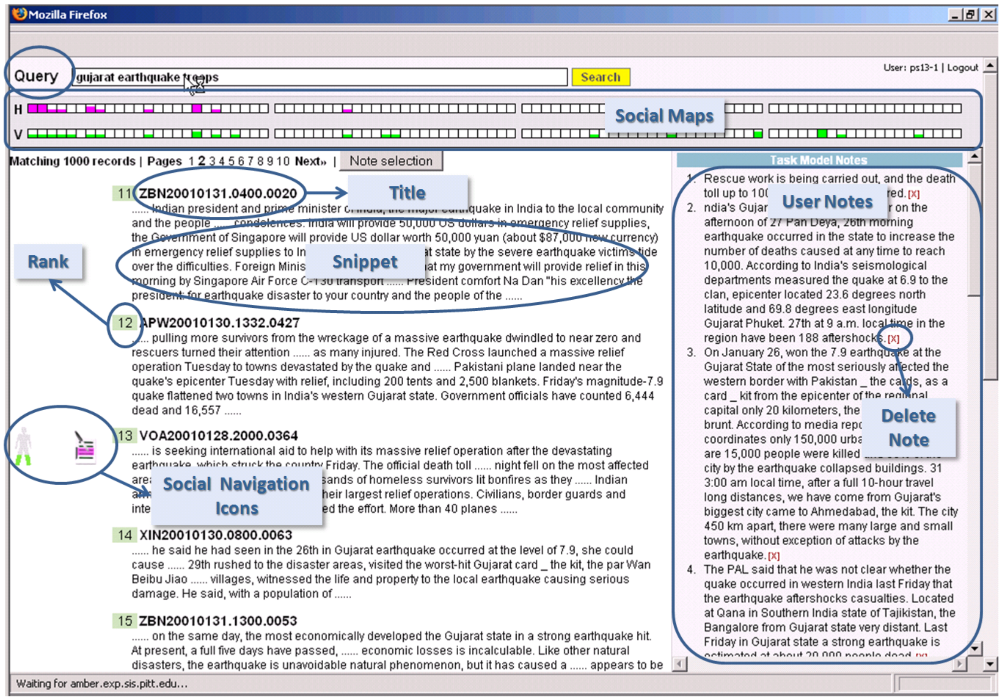


Figure 10: Main interface of the study tool

access the document by clicking on the map cell. The filling level of the cell represents the magnitude of the activity. If the cell is empty, it means the associated document has not been visited or highlighted by anyone. The social maps provide information beyond 10 documents returned on each page of the search results. They were designed to help users have a broader picture of the results in an easy and quick way.

The users can read the full text of each document by clicking on the title of the document. The document was opened in a new window as shown in figure 11. A panel on the right side of the interface shows the list of notes (passages) collected by the current user. To collect notes, users can highlight and save a passage either directly from a snippet shown for each returned document in the list of search results or from the full text of the article, which the user can open by clicking on the document title. The passages saved from this document by past users will be highlighted providing another level of SNS. By default, other users' highlighted passages are shown (in pink). Users can choose to ignore that information and view their own highlighted passages (in yellow).

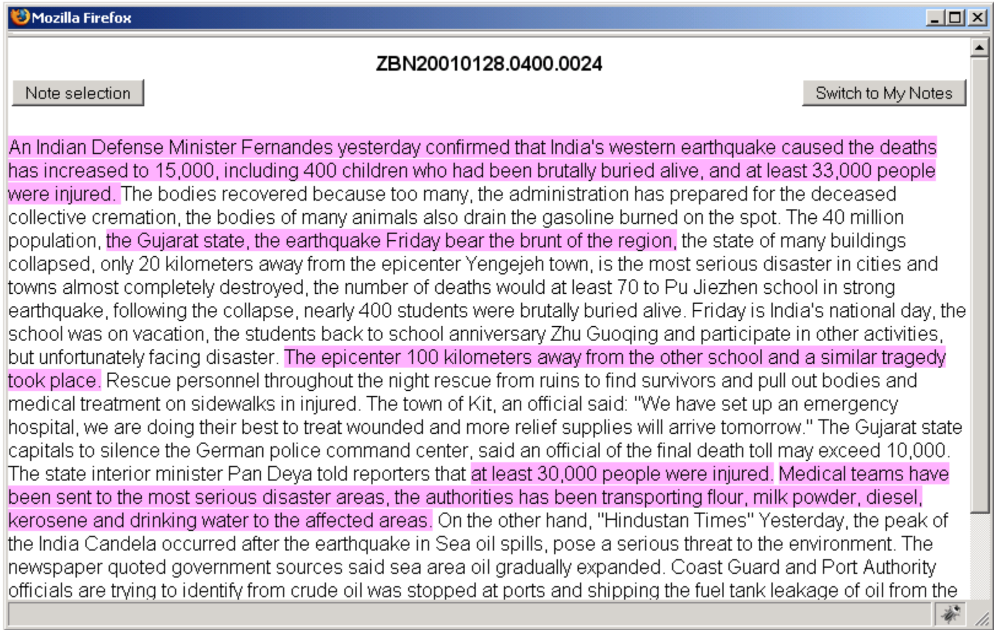


Figure 11: Presentation of the full text of the documents

The system provided SNS through augmenting search results with icons, social maps, and previously highlighted parts of the text. In natural experiments, social navigation cues are generated from the activity of all past users. However, for a controlled study, it is important to make sure that navigation cues stay the same throughout the course of the study and for all participants. To replicate social navigation support while controlling what participants see throughout the study, I used the data from a prior study [2]. For each topic, SNS was offered based on data collected from three prior users who went through the same task working on all 10 questions for the topic. As mentioned earlier, to simulate a collaborative task the questions were divided into two sets. Annotation cues on the interface were based on the note collection activity of prior users excluding questions in which current participants had to work. SNS was not updated with the interaction history of the participants throughout the study. This ensured that all participants had the same opportunity of getting support from social navigation cues.

3.3 DATA COLLECTION

All interactions, of participants with the tool was tracked in a log database. Information in the logs included queries and timestamps of issuing each query, timestamps of every access to a document plus the unique document id, timestamps of closing the document, notes collected, and timestamps of each note collection action.

3.3.1 Eye Movement Data

Eye movements of participants were captured using a Tobii 1570 paired with a 17” LCD monitor (96dpi) set at a resolution of 1024×768. It captures eye position and pupil dilation with a rate of 50 Hz in a natural setting. Eye movement data analysis requires specifying of desired scenes and areas of interest (AOIs) on each scene. Two types of scenes exist in this study, *result* scenes and *text* scenes. I manually assigned the scenes to recordings of 40 sessions (10 participants) using ClearView software. Manual assignment of scenes and AOIs is a very time consuming process and since eye movement data includes details and fine grained data, I limited the analysis of eye-movement data to data collected from half of the participants. 10 participants were selected with the criteria of balancing gender, reading speed, and interpersonal trust level (five participants in each category). The *result* scene is defined when the participants enter a query or look at the search results. The *text* scene is defined when the participants read the full text of the documents.

AOIs were defined manually for each scene using ClearView software. Four AOIs were defined on the *result* scene as shown in figure 12 and two types of AOIs were defined on the *text* scene (figure 13). Icons-AOI and maps-AOI capture attention on social navigation cues, results-AOI captures participants’ attention while assessing the result, and notes-AOI captures any attention on the notes area when the participants review the notes. On the *text* scene htxt-AOI represent the highlighted parts of the text and nhtxt-AOI represent non-highlighted parts of the text. Those AOIs captured participants’ attention on highlighted or non-highlighted text.

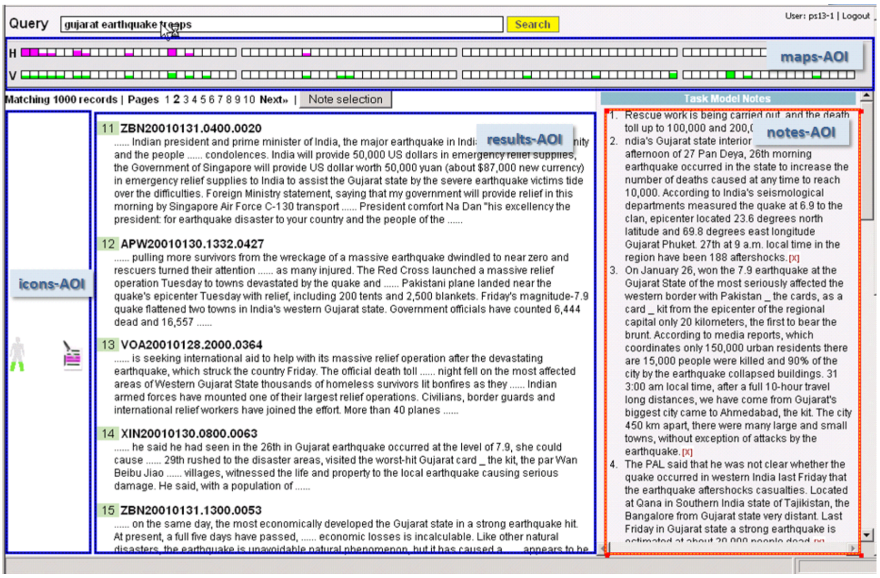


Figure 12: Result scene and AOIs defined on the scene

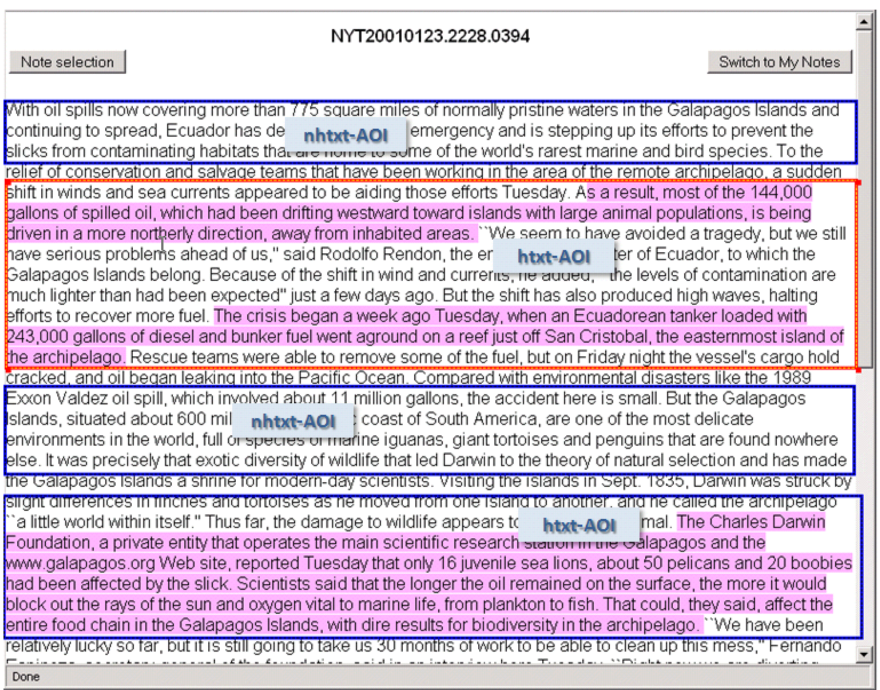


Figure 13: Text scene and AOIs defined on the scene

3.3.2 Interpersonal Trust Classification

Participants were classified as *high trust* and *low trust* based on their responses to an interpersonal trust questionnaire (Appendix B). Numbers 1 to 5 was assigned to responses to each questions (1:Strongly agree, 2:Agree, 3:No opinion, 4:Disagree, 5: Strongly disagree). The first two questions contribute to high trust negatively and questions three, four, and five contribute positively. Overall trust score was calculated as follows:

$$\begin{aligned}&Trust\ Score = \\ &\quad (5 - response(q_1)) + (5 - response(q_2)) + response(q_3) + response(q_4) + response(q_5)\end{aligned}$$

Based on this definition, neutral trust has a trust score of 13 (2+2+3+3+3) and the lower score represents the higher trust. I classified participants with trust scores greater than 13 as *low trust* and the ones with trust scores lower than 13 as *high turst*.

3.4 EXPERIMENTAL DESIGN

The study has a two-by-two design as shown in table 1. It follows a completely random design in which the order of conditions and topics are selected randomly. Each participants went through four topics, two under time pressure and two with SNS selected randomly. Under the *no time pressure* condition the participants had 15 minutes and under the *time pressure* condition they had seven minutes. The time was decided from the prior study. In the prior study participants had 20 minutes to work on a complete set of 10 questions. Therefore, 15 minutes gives enough time to work on five questions enforcing no time pressure. Seven minutes is slightly shorter than enough to work on five questions which enforces time pressure. The *NO-SNS* condition had no social navigation cues. The interface looked the same as figure 10 but with no social navigation icons and no social maps. Also, no indication of prior users' highlights when reading the full article was provided under the *NO-SNS* condition.

Table 1: Study design - conditions

<table>
<tr><td colspan="2"></td><td colspan="2">SNS</td></tr>
<tr><td colspan="2"></td><td>Yes</td><td>No</td></tr>
<tr><td rowspan="2">Time Pressure</td><td>YES</td><td>Topic 1</td><td>Topic 2</td></tr>
<tr><td>No</td><td>Topic 3</td><td>Topic 4</td></tr>
</table>

3.4.1 Topic Selection

As explained earlier, SNS in this study was offered from data collected in a prior study. The previous study included six topics. To minimize the effect of topics, I tried to pick four topics out of six with a similar difficulty level. I used total number of questions, total number of relevant documents in the corpus, average number of relevant documents returned by users, and relevance of collected notes as criteria to match the difficulty level of topics. Total number of questions is the number of questions the user had to respond to for each topic defined as part of GALE topics. Number of questions serves as the basic criterion to make the task difficulty the same. Total number of relevant documents are number of documents in the corpus that at least had one piece of relevant information for one of the questions of the topic. This is available as part of ground truth data. This criterion increases the similarity of chance of retrieving relevant documents for different topics. The average number of relevant documents returned by users is calculated from the prior study data and average relevance is calculated by comparing the the prior study data against ground truth. These two criteria together are a representation of the difficulty of finding relevant information for users for each topic. The criteria and values for each topic is shown in table 2

3.4.2 Question Selection

To simulate a more realistic collaborative task, I divided the questions for each topic into two sets. The first set was used to generate social navigation cues and the second set was assigned

Table 2: Topic selection criteria

	T1	T2	T3	T4
Number of questions	10	10	10	10
Number of relevant documents in the corpus	40	184	162	52
Average number of relevant documents returned by users	6	10	4	4
Average relevance collected notes	0.83	0.73	0.74	0.75

to participants of this study as their task. To make SNS helpful I divided the questions into two clusters with the criteria of decreasing within-cluster similarity and increasing between-cluster similarity using the data from the prior study. That means SNS guides users to relevant documents that include responses to similar questions; however, the highlighted passages were not necessarily responses to questions of the task in hand. The similarity of two questions is calculated based on shared number of documents including response to both questions.

$$Sim(q_i, q_j) = Size((documents\ responding\ q_i) \cap (documents\ responding\ q_j))$$

That means if four documents include a response to both q_1 and q_2, $similarity(q_1, q_2) = 4$. Based on question similarity, cluster similarity was calculated as:

$$Sim(C_l, C_m) = \underset{q \in C_l}{Average}\left(Sim(q, C_m)\right)$$

Where

$$Sim(q, C_m) = \sum_{q' \in C_m} Sim(q, q')$$

10 questions can be divided in two clusters in $C(10, 5) = 252$ ways. Each cluster of questions has a complementing cluster. Excluding the repetitions, we have 126 cases such as $C = (q_2, q_3, q_5, q_7, q_9)$ and $C' = (q_1, q_4, q_6, q_8, q_{10})$ as the complementing cluster. For each case out of 126 cases, I calculated within-cluster similarity ($Sim(C, C)$), and between-cluster

similarity ($Sim(C, C')$). For each topic, I selected the set of questions with the highest between-cluster similarity and lowest within-cluster similarity.

3.5 PROCEDURE

The procedure of the experiment is shown in figure 14. After acquiring a consent form, eye tracking calibration was performed to ensure reasonable precision of the eye tracking device. Next a demographic and skill questionnaire was administered including questions about age, gender, educational level, major, reading speed, and familiarity with Web search. The questionnaire is provided in appendix A. The questionnaire also included five questions measuring interpersonal trust adopted from a questionnaire by Mooradian et al [44]. The interpersonal trust questionnaire is provided in appendix B. Since the task relies heavily on reading speed, a reading speed test was conducted to measure how good of a reader each participant is. The participants were asked to read a piece of text and their reading speed was measured in terms of number of words they read per minute[3]. The reading session was followed by a comprehension test to ensure that the participant read the text. The participants were categorized into four level as shown in table 3.

Next, the task and the interface was explained to each participant in detail including explaining social navigation cues and different part of the interface. To become familiar with the interface and task, all participants went through a training session with no time limit. Each task session started with a pre-questionnaire which was designed to measure background knowledge of the participants about the topic. Next the participants received the task description and started performing the task under a certain amount of time depending on the condition. All topics' descriptions and background knowledge questionnaires are provided in appendix C. The session ended with a survey of the participants' subjective opinions of the tool. The questionnaire included general questions about the usability of the tool and ease of finding information plus usefulness of SNS. Questions related to SNS were omitted from the survey at the end of sessions with no SNS. Additionally, if a NO-SNS

[3] http://www.readingsoft.com/

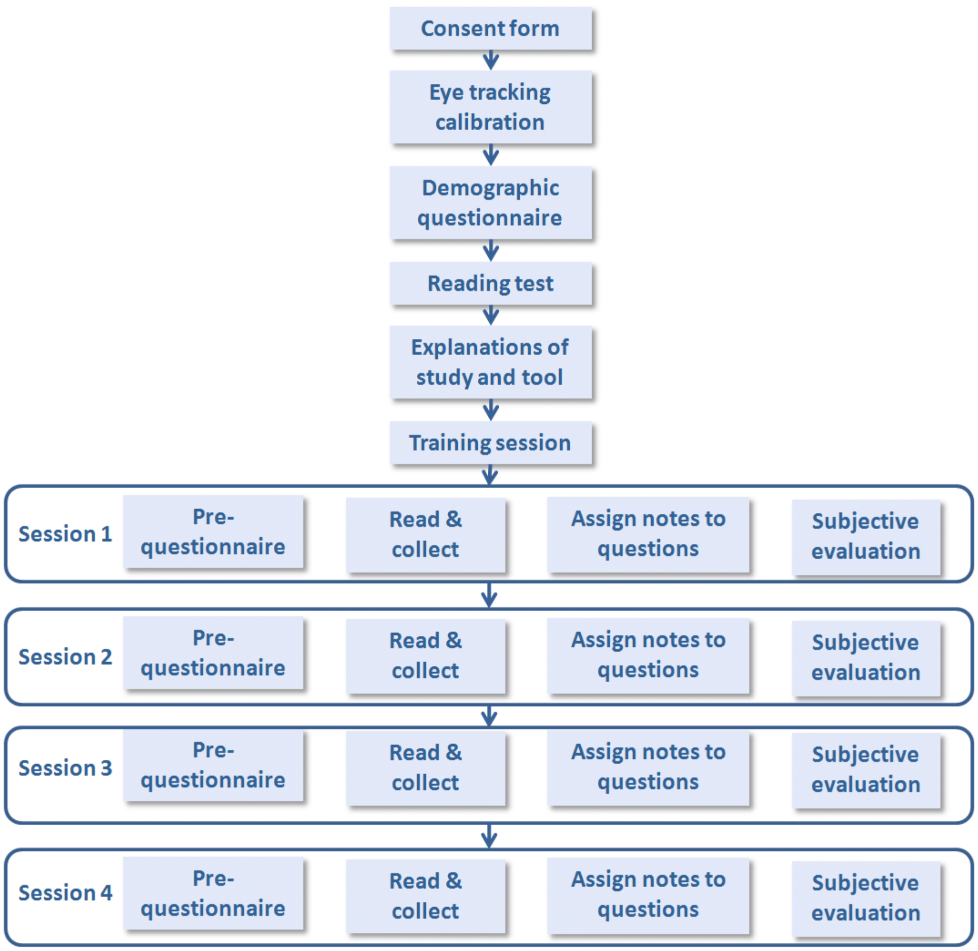


Figure 14: Procedure of the experiment

condition followed a SNS condition, there was one question asking the participant whether it was easier to find information with SNS. Both versions of the questionnaire are provided in appendix D.

Table 3: Categories of readers based number of words read per minute

	Words per minute
Slow readers	< 110
Average readers	110-240
Good readers	240-400
Excellent readers	> 400

3.6 PARTICIPANTS

Twenty participants from students at the University of Pittsburgh from several different disciplines including engineering, information science, life sciences, and humanities were recruited for the study[4]. Participants were paid for their participation in the study. To limit the variability of linguistic abilities I recruited native English speakers. There were nine male and 11 female participants. Their ages ranged from 20 to 35 with the average age equal to 24 (SD=5). All participants were moderately experienced at Web search, reporting that they searched the Web for information several times a week. Except one, participants had not taken any courses in the area of information retrieval. A general demographic of the participants are presented in table 4.

As explained the participants were classified into groups based on their reading speed. Figure 15 shows the distribution of reading speed for all participants and classification of them into the groups. No participants fell into slow reading speed group. Seven were classified as average readers, 11 as good readers, and two as excellent readers.

[4]The study was approved by the Institutional Review Board at University of Pittsburgh and consent forms were collected from all participants

Table 4: Demographic of participants

Gender:	female=11	male=9	
Degree:	Underegrad=14	Master=3	PhD=3
SAT Verbal Score	Max=780	Min=500	μ=620.5, SD=78.4
Reading Speem (wpm)	Max=521	Min=149	μ=267.5, SD=89.8

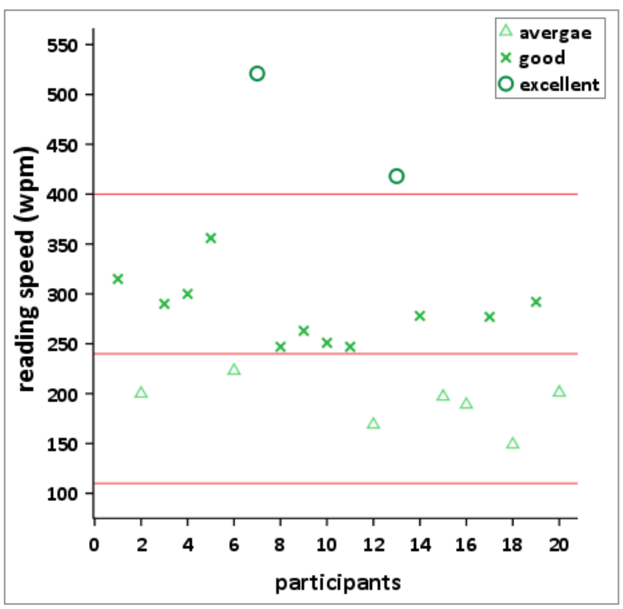


Figure 15: Distribution of participants' reading speed

3.7 DATA ANALYSIS CONSIDERATIONS

Since the study deals with correlated data I performed hypothesis testing with models designed for correlated data. I used the generalized estimating equations (GEE) to fit the model to data and analyze the relationship between the independent variables and dependent factors. GEE is a more flexible statistical tool than the standard general linear model because several types of distribution and different covariance structures of the repeated measures data could be chosen. I fitted two types of models with respect to the distribution of the response variable and goodness of fit: (1) Negative Binomial, and (2) Gamma. The quasi-likelihood under independence model criterion (QIC) was used to select the best fit [46].

The model included session id as the repeated measure within-subjects variable. The dependent factors are the followings:

- SNS: classification factor with two values of 0 and 1 where 1 represents presence of SNS and 0 lack of SNS. The variable is used to study the effect of SNS.
- Time pressure: classification factor with two values of 0 and 1 where 1 represents enforcement of time pressure and 0 no enforcement of time pressure. The variable is used to study the effect of time pressure.
- Gender: classification factor with 1 for female and 0 for male. The variable is used to study the effect of gender.
- Trust: classification factor with two values of 0 and 1 where 1 represents high interpersonal trust and 0 low interpersonal trust. The variable is used to assess the effect of interpersonal trust.
- Reading speed: classification factor with three values of 1,2, and 3 where 1 represent average reading speed, 2 good, and 3 excellent. The variable is used to assess the effect of interpersonal trust.
- Topic: classification factor with four values of 1,2,3, and 4 representing four different topics in the study. The variable is used to study any topic effect.
- Order: classification factor with four level of 1,2,3, and 4 representing four sessions. The variable is used to assess any order effect.

Additionally the model includes the following interactions: sns*tp, sns*gender, sns*trust, sns*reading speed, tp*trust, tp*gender, and tp*reading speed.

Before analyzing the data, I assessed the correlation between factors; specifically, correlation of gender and reading speed and correlation of gender and trust was evaluated. The analysis shows no significant correlation of gender and trust or gender and reading speed. Spearman's correlations coefficients are presented in table 5. None of the factors were excluded from the data analysis.

Table 5: Correlation of Gender and Reading Speed, and Trust

		trust	reading speed
gender	ρ	-.161	.302
	p-value	.468	.189

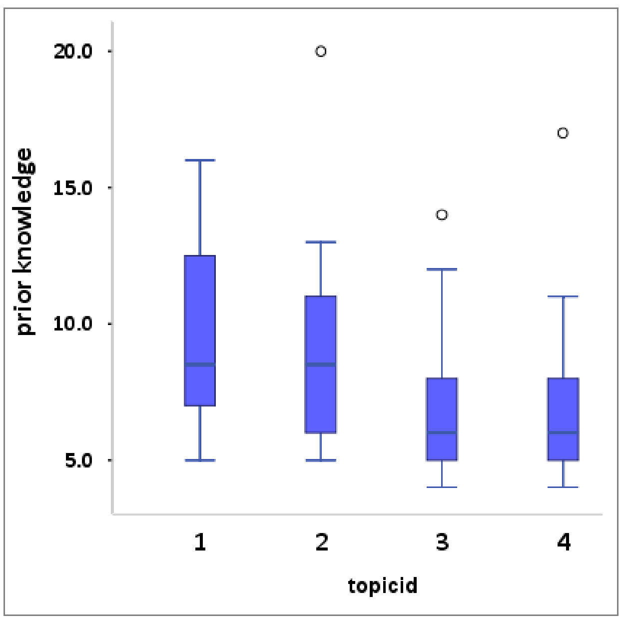


Figure 16: Range of participants' prior knowledge of different topics.

Additionally, participants' prior knowledge was assessed to ensure homogeneity of background knowledge. As explained in the procedure, a background knowledge questionnaire was administered at the beginning of each session (appendix C). The questionnaire included four questions asking participants whether they had already read any news about the topic, how much they already knew about the story or the location the story happened, and what their general knowledge of the topic was. The possible responses to each question ranged from 1 (no knowledge at all) to 10 (expert); i.e. the maximum score they could get for each topic was 40. Figure 16 shows the range of participants' prior knowledge on different topics. The data shows that all participants had limited background knowledge of all topics. There is only one outlier who had 50% knowledge of the second topic. Further investigation of that participant's response shows that he had good knowledge of the general field related to the second topic which is the least contributing one to the task. Since the overall score was still only 50%, I made the decision to not treat him as an outlier and include all the participants in the data analysis.

4.0 RESULTS

The analysis of data is divided into three main sections: (1) Effect of SNS on users' search behavior; (2) Effect of SNS on users' performance of the task; (3) Subjective opinion of users about the task difficulty and SNS under different conditions. User action logs and eye movement data were used to study the first two parts and the third part was evaluated utilizing participants' responses to the subjective questionnaire.

4.1 EFFECT OF TOPIC AND ORDER

I employed completely random design to avoid the effect of order; however, throughout the analysis, the effect of order was checked and any significant effect was reported.

As explained earlier, I used data from the prior study to select four topics comparable with respect to the level of difficulty; however, the result of the current study shows a significant effect of topics on some of the variables. Overall, several measures show that topic 2 and 4 were easier than topic 1 and 3. It is important to mention that the chance of seeing documents with SNS is not the same for all topics and depends on participants' queries and topics. I define the chance of seeing an augment document as:

$$Chance\ of\ seeing\ an\ augmented\ document = \frac{\#\ of\ seen\ augmented\ documents}{Total\ \#\ of\ seen\ documents}$$

Figure 17 presents the chance of seeing documents augmented with social navigation icons for each topic. One-way ANOVA comparison of the means with Bonoferroni post hoc comparisons shows that the chance is significantly higher for topic 2 and 4 comparing with topic 3(t2 vs. t3: Mean diff=.28, SE=.072, p-value=.003; t4 vs. t3: Mean diff=.33, SE=.07, p-value=.000)[1]. This might have played a role in making the second and fourth topics easier for the participants. Throughout the analysis any significant effect of the topics is reported.

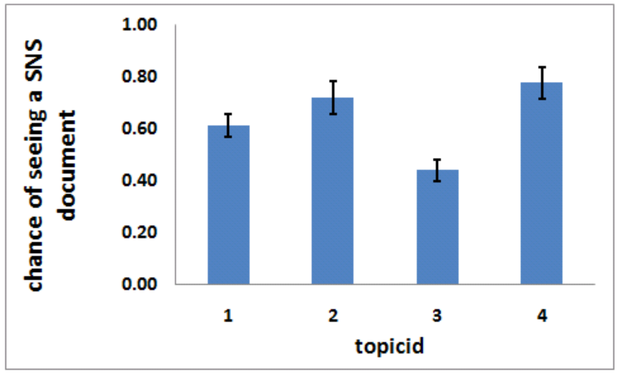


Figure 17: Chance of seeing a SNS-document

4.2 EFFECT OF SNS ON USERS' SEARCH BEHAVIOR

SNS can affect users' search behaviors in terms of navigation decisions and search effort. If participants' navigation decisions are affected by SNS, they are more likely to access documents augmented with social navigation icons, access documents from social maps, and be influenced by highlighted text. The effect can be magnified under time pressure. As suggested by the literature, female participants are expected to follow SNS more. Additionally I expected to observe an effect on interpersonal trust that participants classified in the *high trust* group follow SNS more.

Any navigation support is expected to reduce users' efforts for finding information, especially for those who utilize the support. In the current experiment, the information search effort can be measured by number of queries, length of queries, collecting note rate, time spent searching, and average fixation duration on search results. Larger number of queries with larger lengths can be associated with more effort and more difficulty in finding the

[1] ANOVA assumptions of normality of distribution and homogeneity of variance were met (Kolmogrov-Smirnov test of normality: stat=.119, df=40, p-value=.161; Levene test of homogeneity of variance: stat=.75, df=(3,36), p-value=.529)

desired information. Higher collecting note rate, which means accessing more relevant versus irrelevant documents and less time spent searching, are indicative of less effort. Longer fixations on results can be associated with higher effort since it is an indication of more information processing that can be due to higher density of information or more difficult information. I expected to observe less effort in terms of these measures as an effect of SNS especially for participants who follow SNS the most.

I define an **SNS-document** as a document where the link to it is augmented with either visiting or annotating icons.

4.2.1 Measurements

- **SNS-documents access rate**: proportion of the number of accesses to SNS-documents to total number of accesses normalized by chance of seeing an SNS-document. It is defined as:

$$SNS - documents\ access\ rate = \frac{\frac{size(access(SNSdocuments))}{size(access(all\ documents))}}{chance\ of\ seeing\ an\ SNSdocument}$$

- **SNS-documents collecting note rate**: proportion of the number of notes collected from SNS-documents to the total number of collected notes. It is defined as:

$$SNSdocuments\ collecting\ note\ rate = \frac{\frac{size(notes(SNSdocuments))}{size(notes(all))}}{chance\ of\ seeing\ an\ SNSdocument}$$

- **Maps access rate**: proportion of the number of accesses from the social maps to the total number of accesses. It is defined as:

$$Map\ access\ rate = \frac{size(access(from\ maps))}{size(access(all\ documents))}$$

- **Collecting note rate**: proportion of the number of documents the user accessed and collected notes from to the total number of accessed documents. It is defined as:

$$CollectingNoteRate = \frac{size(access(documents) \cap note(documents)))}{size(access(all\ documents))}$$

4.2.2 Following Social Navigation Cues: Detailed Results

Following of social navigation was measured by analyzing SNS-documents access rate, SNS-documents collecting note rate, SNS maps access rate, and eye tracking measures of fixation count on social navigation icons, social maps, and highlighted passages of text.

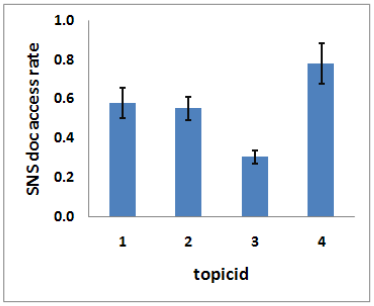
SNS-documents access rate: SNS-documents access rate captures the effect of SNS on participants' document access behaviors. SNS-documents access rate was calculated for each session with SNS (two sessions out of four). The result of analysis is shown in table 6. There is a significant effect of topic, time pressure, and significant interaction of time pressure and reading speed. Pairwise comparison of topics with Bonferroni adjustment shows significant differences of topic 3 with topic 1, 2 and 4. Participants followed SNS significantly less for topic 3 as shown in figure 18(a). As explained earlier, the chance of seeing an SNS-document was significantly lower for topic 3 which might have contributed to this difference as well, even though the measure was normalized.

The significant effect of time pressure shows that participants follow SNS significantly more under time pressure (figure 18(b)). The significant interaction of time pressure and reading speed shows that time pressure has significantly more effect on participants with excellent reading speed (figure 18(c)). Participants with weaker reading abilities follow SNS even when there is no time pressure, but the need of navigation support for participants with strong reading abilites changes by time pressure.

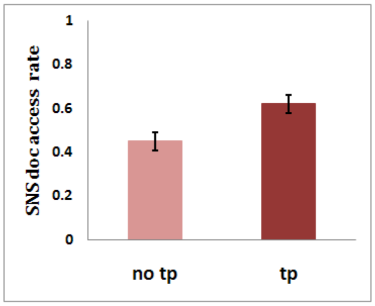
Table 6: Analysis of SNS-documents access rate

	Wald χ^2	df	p-value
time pressure	16.90	1	<.001
topic	34.59	3	<.001
time pressure*reading speed	8.72	2	.01

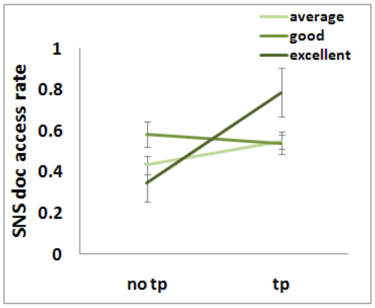
SNS-documents collecting note rate: While SNS-documents access rate presents the effect of SNS on participants' document access behaviors, SNS-documents collecting note rate measures the effect beyond just accessing those documents and represents whether the participants note collection behavior was affected by SNS. This measure was also calculated



(a) Effect of topic



(b) Effect of time pressure



(c) Interaction of time pressure and reading speed

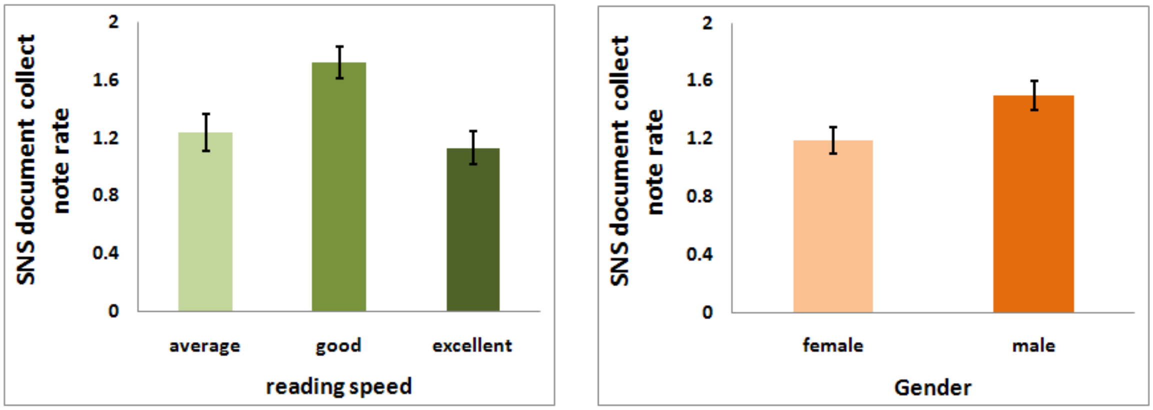
Figure 18: SNS-documents access rate

for two sessions with SNS and it takes into account the chance of seeing an SNS-document. The result of the analysis is shown in table 7. There is a significant effect of gender and reading speed. Male participants were significantly more likely to collect notes from SNS-documents(figure 19(b)). Pairwise comparison of reading speed shows that good readers collected significantly more notes from the SNS-documents. The result is shown in figure 19(a).

Table 7: Analysis of SNS-documents collecting note rate

	Wald χ^2	df	p-value
gender	5.53	1	.019
reading speed	21.12	2	<.001

SNS maps access rate: SNS map access rate measures the effect of SNS on participants' search behaviors from a different perspective. Most Web users are familiar with search results presented as an ordered list of titles with a short snippet. Taking into consideration social navigation icons, it still matches the traditional searching style. It can serve as an additional hint while reviewing the result. However, using SNS maps falls into a different searching behavior which is primarily based on the activities of other users. SNS map access rate was also calculated for SNS sessions. Contrary to usage of social navigation icons,



(a) Effect of reading speed

(b) Effect of gender

Figure 19: SNS-documents collecting note rate

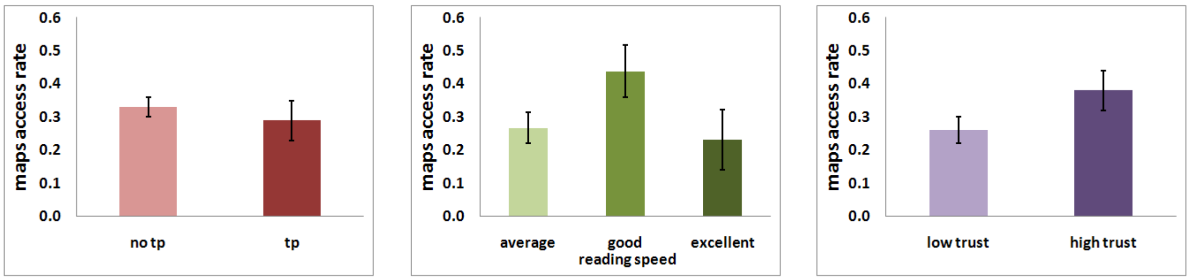
the result shows no significant effect of time pressure on usage of the maps. Participants used the maps non-significantly more under no time pressure (figure 20(a)). This can be a result of the fact that under time pressure the participants had very limited time to explore documents. The data shows that, on average, they accessed 3.2 documents with an average rank of 1.5 (median=1) and maximum rank of 10. That means under time pressure, the majority of participants did not go beyond rank three, and no participants considered any document beyond the first page. It is also possible that participants under time pressure have a preference of relying on the most familiar strategy, which is making a decision based on ranking, title, snippet, while getting the additional support from SNS icons.

There is a significant effect of trust and reading speed on the usage of maps as shown in table 8. Participants with high trust used the maps significantly more (figure 20(c)). The result matches the utility of the maps. As mentioned above, the decision to access a document from the maps is made primarily based on the activities of others. Therefore, participants who have higher interpersonal trust are expected to be more likely to access a document from the maps. Similar to SNS-documents access rate, map access rate is significantly higher for readers with good reading speed (figure 20(b)).

Fixation count of social navigation icons and social maps: Fixation count on

Table 8: Analysis of social map access rate

	Wald χ^2	df	p-value
trust	4.56	1	.033
reading speed	8.29	2	$<.016$



(a) Effect of reading speed (b) Effect of reading speed (c) Effect of trust

Figure 20: Social maps access rate

the icons AOI is an indication of the participants' attention to the icons. This is a stronger evidence of paying attention to icons in comparison with counting accesses. Access to a document could have happened due to different reasons such as the rank, the title, or the relevancy of the snippet. I assessed the effect of time pressure, gender, trust, and reading speed on fixation count on social navigation icons. To normalize the number of fixations for long and short sessions, I calculated normalized fixation count as follows:

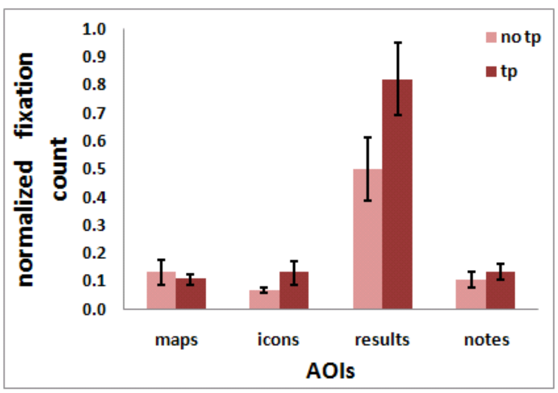
$$Normalized\ fc = \frac{fc(AOI_i)}{\displaystyle\sum_{AOIs\ on\ result\ scene} fc(AOI)}$$

Since the goal is analysis of fixation count on social navigation cues, this analysis includes only sessions with SNS and the result scene with four AOIs as shown in figure 12. The significant reuslts of pairwise comparisons are shown in table 9.

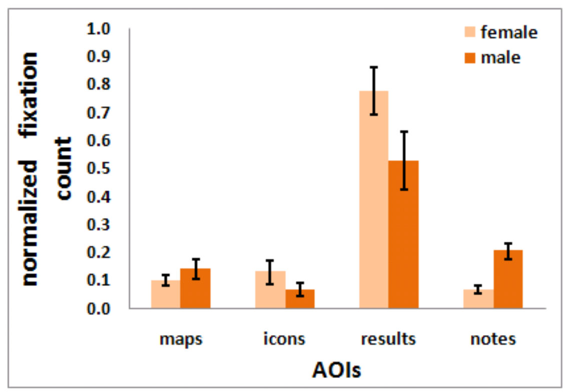
The effect of time pressure is shown in figure 21(a). Pairwise comparison analysis shows no significant difference in terms of fixation count on maps; however, under time pressure

Table 9: Analysis of fixation count on result scene

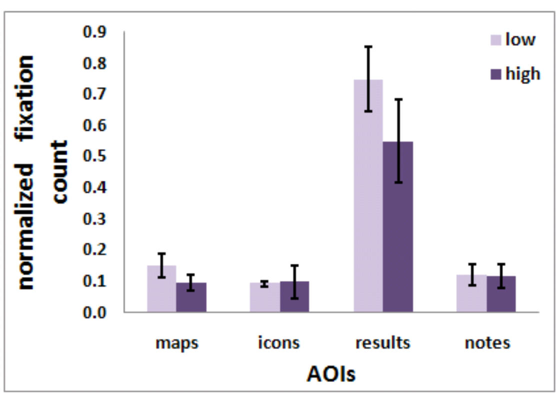
AOI	Effect	Mean diff	SE	df	p-value
icons	tp vs. no-tp	-.09	.048	1	.05
results	tp vs. no-tp	-.56	.17	1	.001
results	female vs. male	.35	.17	1	.037
results	low vs. high trust	.62	.25	1	.015
maps	average vs. good readers	-.09	.05	1	.05



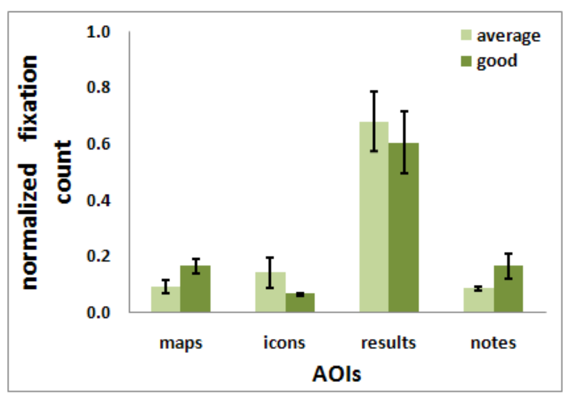
(a) Effect of time pressure



(b) Effect of gender



(c) Effect of interpersonal trust



(d) Effect of reading speed

Figure 21: Normalized fixation count on result scene AOIs

there are significantly higher number of fixations on social navigation icons which matches the result of higher SNS-documents rate under time pressure. Also, there is a significant difference in terms of number of fixations on the results section under time pressure.

The effect of gender is shown in figure 21(b). Female participants have non-significantly higher number of fixations on social navigation icons while male participants have non-significantly higher number of fixations on social maps. Female participants have significantly higher fixations on results AOI. The result suggests that female participants paid more attention on the results including the icons next to them while male participants explored other parts such as notes and maps.

The effect of trust is shown in figure 21(c). There is no significant effect of trust on number of fixations on maps and icons. Participants with low trust have a significantly higher number of fixations on results AOI. The results suggest that both group have examined social navigation cues but participants with high trust chose to make more use of social maps.

The effect of reading speed is shown in figure 21(d). Participants with good reading speed have a significantly higher number of fixations on social maps which is consistent with their higher access to SNS-documents. Average readers might have not had enough time to explore the results beyond the list of results. As a result they did not pay attention to social maps which can be the cause of not using the maps too.

Order Number of Fixations on Results: In addition to fixation count, I analyzed the order number of fixations on all AOIs in the *result* scene to assess whether there is any pattern in terms of the order the participants reviewed different AOIs. The result is shown in figure 22. The result shows that participants first looked at icons, then results, and significantly later they looked at the maps. Under time pressure, participants looked at icons slightly before the results which suggest they reviewed the icons prior to making a decision to access a document from the list. However, under time pressure, the result suggests that participants were going back to icons after reviewing results creating a wider range of order number of looking at icons.

Fixation count of highlighted text: As mentioned earlier, in the SNS sessions, while reading the full text, participants could see passages highlighted by prior users. Fixation count on the highlighted passages represents participants' attention on those parts of the

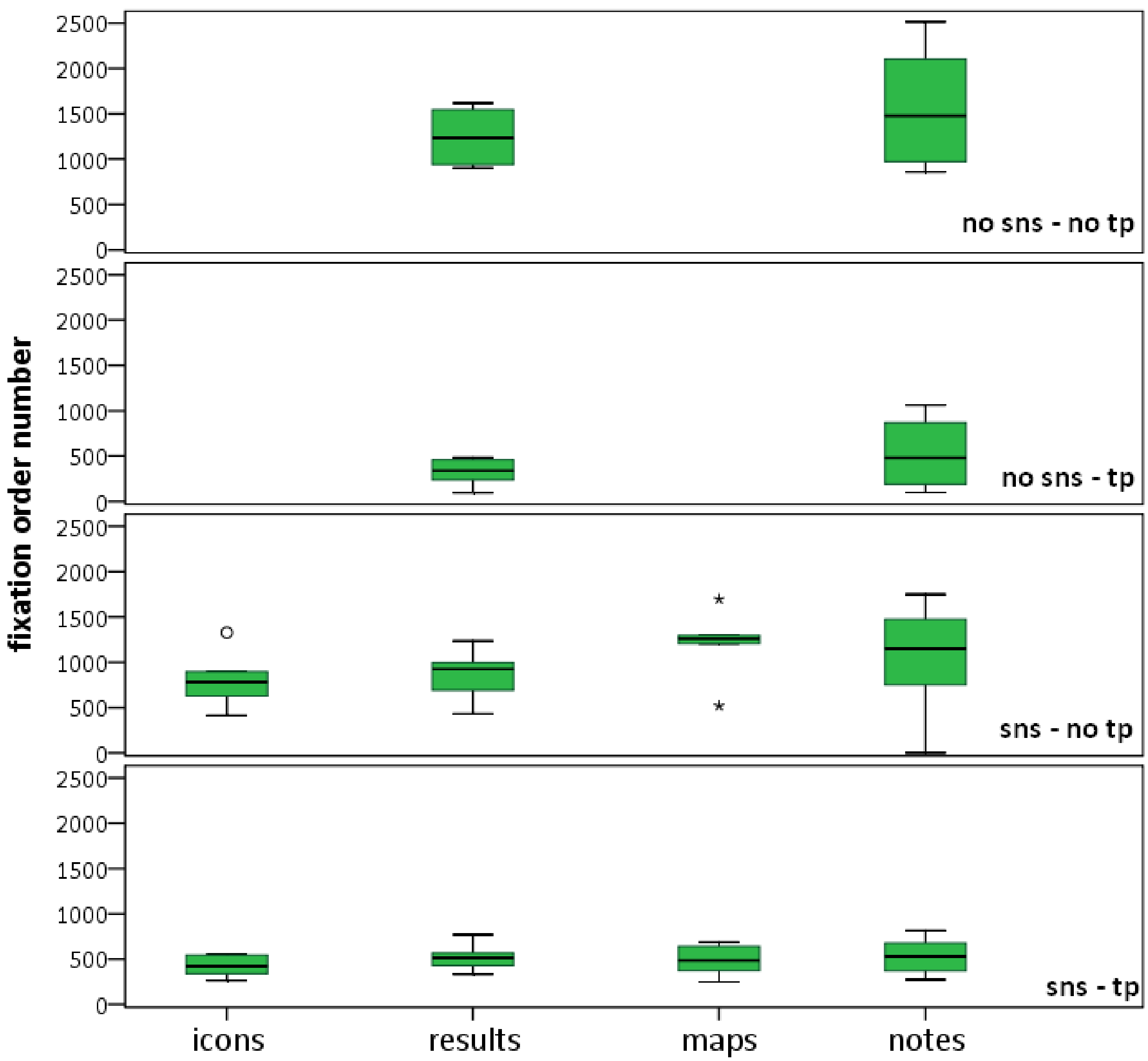


Figure 22: Order number of fixation on AOIs in *result* scene

text. On average 73.5% of fixation on the text area happened on non-highlighted text, and 26.5% on highlighted text. However, the proportion of the highlighted part and the non-highlighted part of the text was not the same. Therefore, I normalized the fixation count on each document by dividing the number of fixations by the following normalizing factor:

$$normalizing\ factor = \frac{length(highlighted\ text)}{length(document)}$$

Length is measured in number of characters. The average normalized fixation count is shown in table 10. Comparison of the mean of the normalized fixation count on highlighted and non-highlighted text for each session shows no significant difference. As mentioned earlier, the highlighted text was not entirely related to questions the participants were working on; instead there were responses to similar questions. This result suggests that participants did pay attention to non-highlighted parts of the text as much as highlighted parts. It means SNS helped the participants to get to the relevant documents, but within a document they relied on their own judgments.

Table 10: Average normalized fixation count on highlighted and non-highlighted text

	Mean	SD	SE
highlighted text	1.52	1.26	.32
non-highlighted text	.97	.23	.06
	Z	2-tailed p-value	
Wilcoxon Signed Ranks test[2]	-1.03	.301	

Moreover, I analyzed the data to assess whether there was any effect of condition and personal factors on attention to highlighted text. The result of analysis is shown in table 12. There is a significant effect on order, topic, and significant interactions of time pressure and gender, trust, and reading speed. The first and last session had a smaller number of fixations on highlighted text (table 11). Novelty might have been the cause of higher fixations for the first session and exhaustion might have been the cause of higher fixations in the last session. Participants had significantly higher number of fixations on highlighted text for the second topic. As mentioned earlier, the chance of seeing an SNS-document was higher for second topic. That could have contributed to this difference as well.

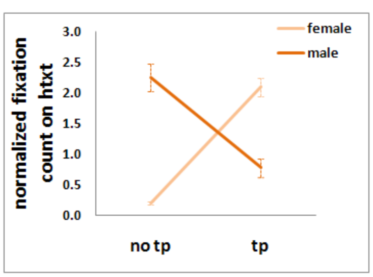
Table 11: Average fixation count on highlighted and non-highlighted text

	1	2	3	4
order	.72(SE=.14)	1.26 (SE=.11)	1.14 (SE=.09)	.74 (SE=.06)
topic	.50 (SE=.19)	2.43 (SE=.18)	.80 (SE=.10)	.78 (SE=.03)

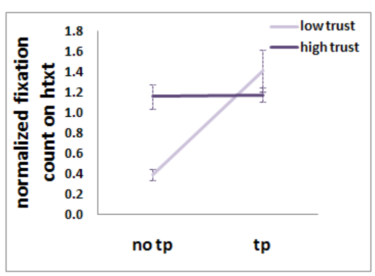
Table 12: Analysis of normalized fixation count on highlighted and non-highlighted text

	Wald χ^2	df	p-value
order	36.48	2	<.001
topic	207.49	3	<.001
time pressure*gender	114.87	1	<.001
time pressure*trust	53.28	1	<.001
time pressure*reading speed	66.33	1	<.001

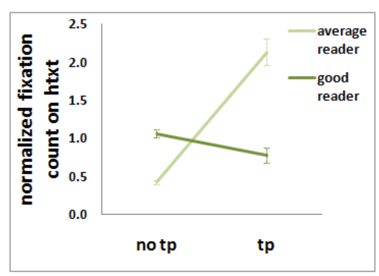
The interaction of time pressure and personal factors (gender, trust, and reading speed) is presented in figure 23(a), 23(b), and 23(c). Female participants have a significantly higher number of fixations on highlighted passages under no time pressure; however, under time pressure it drops significantly, while the number of fixations on those passages increases significantly for male participants. This is a rather surprising result which is difficult to explain given the rest of findings of the study.



(a) time pressure and gender



(b) time pressure and trust



(c) time pressure and reading speed

Figure 23: Interaction of time pressure and personal factors on fixation count on highlighted text

Participant with high trust were not influenced by time pressure, but the low trust group had a significantly higher number of fixations on highlighted passages. That means the low trust group had mostly ignored the activity of others at the text level as long as they had enough time to read the text. Time pressure had affected the average readers significantly more. As expected, under time pressure, the average readers had a significantly higher number of fixations on highlighted passages.

Even though the participants paid similar attention to highlighted and non-highlighted parts of the text, the data about the order number of fixations shows that their attention was drawn to highlighted text initially after opening the document (figure 24) and the difference is significant (Wald χ^2=22, df=1, p-value<.001). That means the participants evaluated the highlighted passages first and, since it did not satisfy their information need, they started reading the rest of the text.

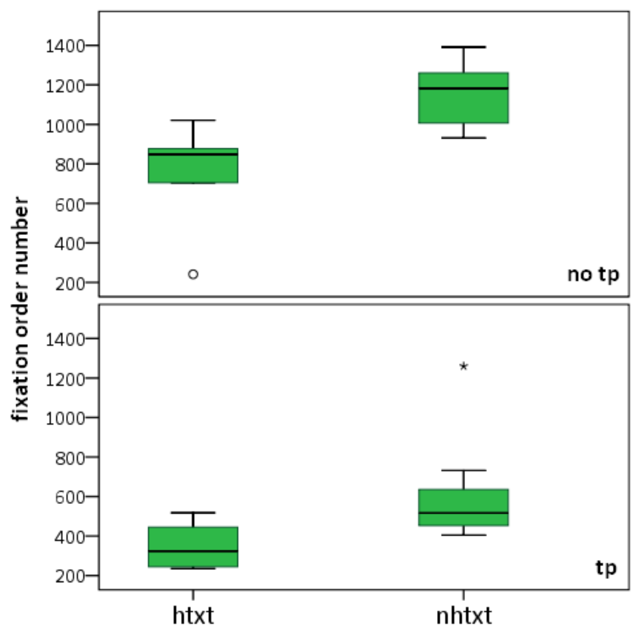


Figure 24: Order number of fixation on highlighted and non-highlighted passages

As mentioned earlier, the participants had an option to turn off highlighting of passages. Overall, five participants out of 20 switched off the highlights for seven trials out of 40. Information about the participants and condition is shown in table 13. Due to low number of participants who choose to turn off highlighting, no strong claim can be made here. The data suggests a pattern that female and low trust participants may be more likely to make that decision and it is more likely to happen under no time pressure.

Table 13: Switching off highlighting

userid	condition	gender	trust	reading speed
user1	no-tp	male	high	good
user2	tp	female	high	average
user3	tp and no-tp	male	low	excellent
user4	no tp	female	low	good
user5	tp and no-tp	female	low	good

4.2.3 Effort

The following measures were used to assess the effort for performing the task:

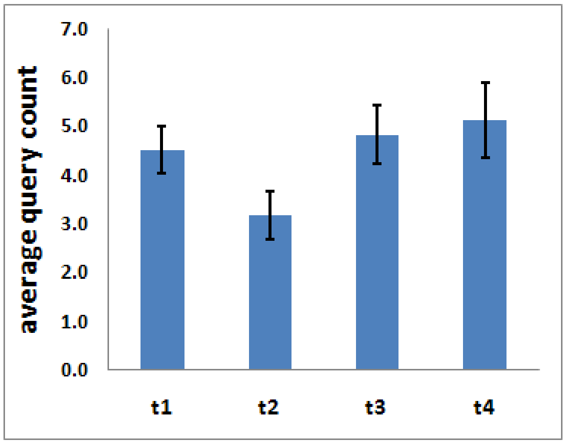
- Number and length of queries
- Collecting note rate
- Time spent searching (TSS) versus time spent reading (TSR)
- Fixation duration on search results

Number of Queries: SNS highlights potentially relevant documents for each query; therefore, it is possible for users to find their desired information by trying lower number of queries.

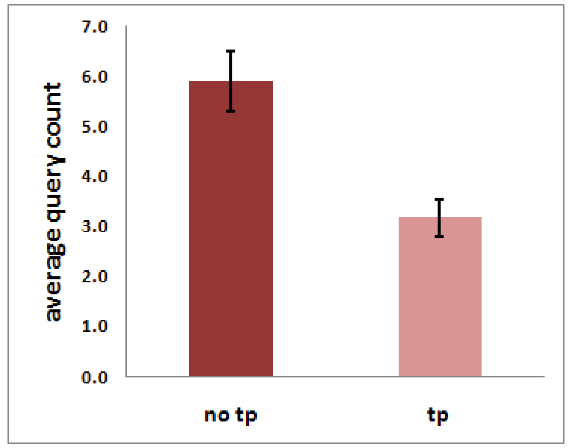
Table 14: Analysis of number of queries

	Wald χ^2	df	p-value
time pressure	46.195	1	<.001
topic	16.776	3	.001
sns*trust	3.629	1	.057
sns*reading speed	9.253	2	.010

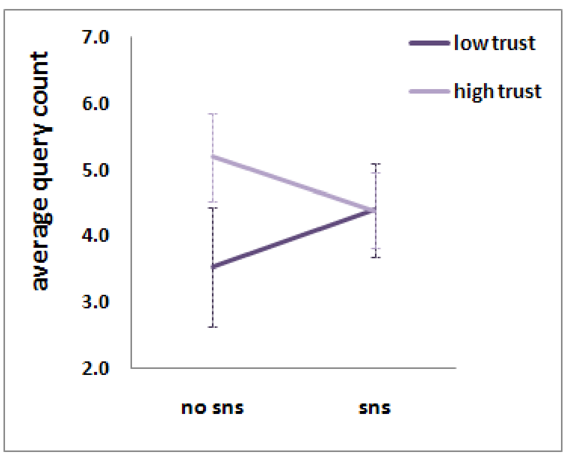
The result of analysis of the number of queries is shown in table 14. The significant effect of time pressure is an expected result. Participants had double the time under no time pressure and they entered about twice more queries as shown in figure 25(b). There is a



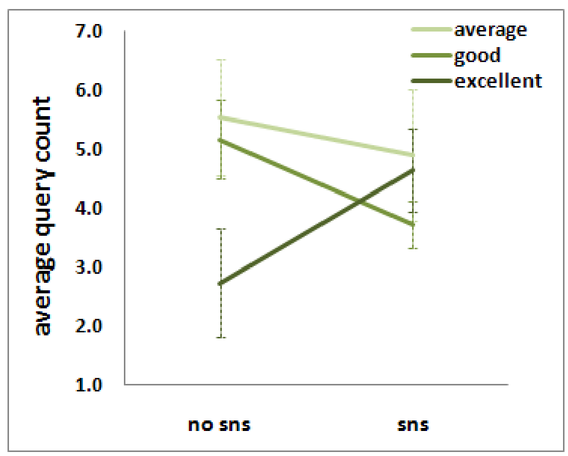
(a) Effect of topic



(b) Effect of time pressure



(c) Interaction of SNS and trust



(d) Interaction of SNS and reading speed

Figure 25: Average query count

significant effect of topics which shows that participants entered significantly lower number of queries for topic 2 comparing with topic 3 and 4 (figure 25(a)). This is another indication that subject had less difficulty finding information for the second topic. There is a significant interaction of SNS and trust, and a significant interaction of SNS and reading speed. Interaction of SNS and trust (figure 25(c)) shows that participants with high trust entered significantly fewer queries under the SNS condition. The interaction of SNS and reading speed (figure 25(d)) shows that participants with good reading speed entered a significantly smaller number of queries under the SNS condition. Result in previous section showed that participants with high trust and good reading speed followed social navigation cues significantly more. Consistent with that result, here the interactions suggest that participants who utilized social navigation cues more, spent less effort in terms of number of queries.

Query Length: Query length is calculated as an average number of words per query for all queries of each session. The result of analysis (table 15) shows significant effect of topic and significant interactions of SNS*time pressure, SNS*gender, and SNS*trust. The interaction of SNS*time pressure (figure 27(a)) suggests that, under time pressure, the average length of queries decreases significantly when SNS is provided.

Previous studies of gender differences in Web search suggest that females are more likely to use longer queries [19]. The significant interaction of gender and SNS as shown in figure 27(b) shows similar result that under the no-SNS condition female participant use longer queries; however, SNS eliminates the difference. A similar result is observed as an interaction of SNS and trust (Figure 27(c)). Participants with high interpersonal trust use shorter queries when SNS is provided. Additionally, there is a significant effect of the topic suggesting that participants entered significantly shorter queries for topic 4 as shown in figure 26. Again, I remind that the chance of seeing SNS-documents is higher than topic 1 and 3 and similar patterns is observed in figure 26. Overall the result suggest that SNS decreases participants' efforts in term of query length especially under time pressure and for those who rely on SNS more.

Collecting Note Rate: The average collecting note rate is .62 (SD=.20) which means participants on average collected at least one note from 62% of the documents they had accessed. SNS can guide users to more relevant documents. As a result, I expected to

Table 15: Analysis of length of queries

	Wald χ^2	df	p-value
topic	77.592	3	$<.001$
sns*time pressure	9.486	1	.002
sns*gender	8.633	1	.003
sns*trust	5.144	1	.023

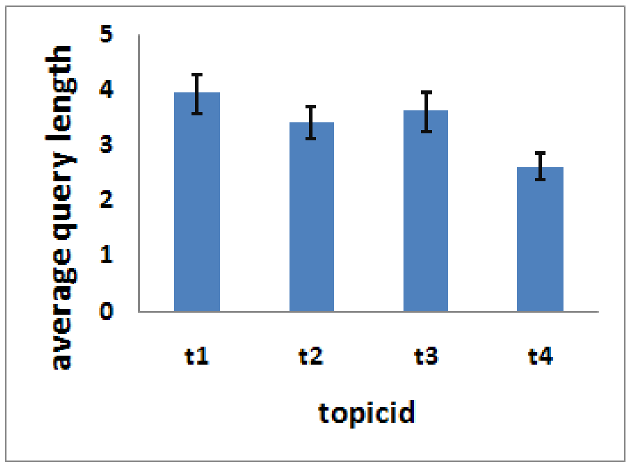
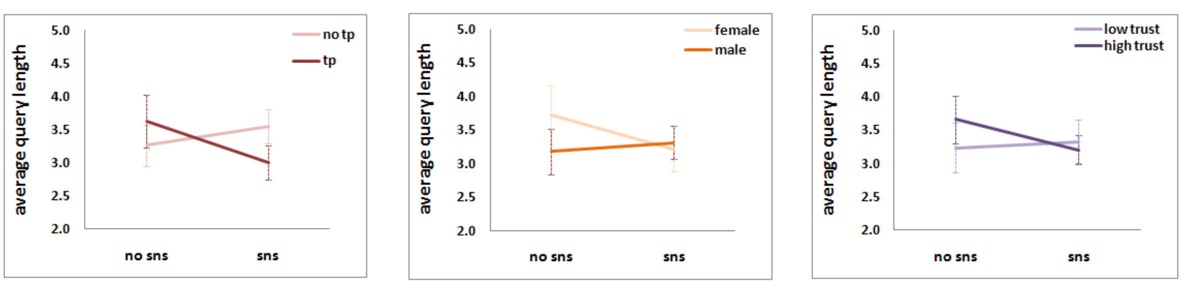


Figure 26: Effet of topic on average query length (in words)



(a) SNS and time pressure (b) SNS and gender (c) SNS and trust

Figure 27: Average query length (in words)

observe a higher collecting note rate while SNS was provided; however, the analysis shows no significant effect of SNS. There is significant effect of time pressure and reading speed. Collecting note rate was significantly higher under time pressure. As mentioned earlier, under time pressure, participants opened very few document form high ranks. Therefore, it is expected to achieve a significantly higher collecting note rate under time pressure. Participants with excellent reading speed had a significantly higher collecting note rate as shown in table 16. This is also an expected result. Excellent readers are faster in reading the snippets; therefore, they are more successful in selecting more relevant documents.

Table 16: Average collecting note rate for different reading speed

		μ	SE	Wald χ^2	df	p-value
time pressure	yes	.71	.04	8.05	1	.005
	no	.57	.03			
reading speed	average	.60	.02	11.81	2	.003
	good	.61	.03			
	excellent	.71	.03			

Time Spent Searching (TSS): The task involves two main steps: Searching for relevant documents and collecting notes from relevant documents. I define TSS as the time spent on searching to find a desirable document and TSR as the time spent reading the document and collecting notes. SNS is designed to guide users to relevant documents faster; therefore, I expected that when users are supported with social navigation cues they spend less time searching and more time reading and collecting notes. To assess the effect of SNS on TSS, for each session I calculated total time the user had spent on the result pages and total time they spent on the text pages. I define TSS% as follows:

$$TSS\% = \frac{TSS}{TSS + TSR}$$

The result of the analysis is shown in table 17. Contrary to my expectation, there is no significant effect of SNS. There is significant effect of order, topic, time pressure, and reading speed. Participants spent significantly less time searching on the last two sessions which

Table 17: Analysis of TSS

		μ	SE	Wald χ^2	df	p-value
order	1	22.02	3.19	11.786	3	.002
	2	22.54	3.10			
	3	16.27	1.83			
	4	16.70	2.46)			
topic	1	22.39	2.10	9.367	3	.003
	2	11.75	1.83			
	3	22.24	2.69			
	4	23.06	3.72)			
time pressure	yes	14.68	1.54	10.527	1	<.001
	no	25.02	3.27			
reading speed	average	23.69	3.63	7.63	2	.022
	good	21.34	2.97			
	excellent	13.92	3.05			

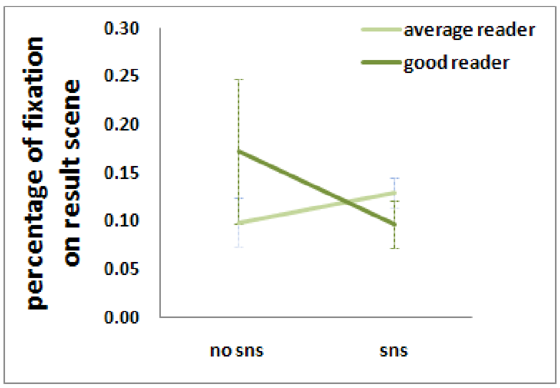
might be due to being tired. Under time pressure participants spent significantly less time searching which means they preferred to spend their time collecting notes and they did less exploration of the results. Participants with excellent reading speed spent significantly less time on searching. This is quite a surprising result. It is expected that participants with excellent reading speed spent less time on reading and more time on searching and exploring results. However, the result from the previous section shows that they have a significantly better collecting note rate. This suggests that they are faster in reading the snippets and evaluating the relevancy of documents. Significant effect of topic shows that participants spent significantly less time on searching when working on topic 2.

Time spent is frequently treated as reliable measure of users' attention and interest[10]. At the same time, it is often criticized that there is a lot of noise associated with time spent data. Users might have been distracted from the main task for different reasons. Even though time is more reliable under a controlled experiment, in presence of eye tracking, fixation count can shed more light into true actions of users. Therefore, to compare users' searching versus reading activities, I analyzed the percentage of fixations on the *result* AOI versus the *text* AOIs. The result is shown in table 18. Similar to TSS data, there is a significant effect of topic, and time pressure. Participants have a lower percentage of fixations under time pressure and a lower percentage of fixations on topic 2. However, there is no significant effect of reading speed or order as suggested by time spent searching.

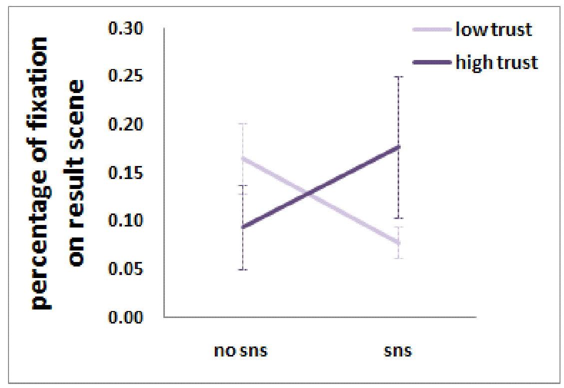
Table 18: Analysis of fixation counts on *result* scene

		μ	SE	Wald χ^2	df	p-value
topic	1	.18	.04	70.19	3	<.001
	2	.05	.01			
	3	.11	.02			
	4	.20	.06			
time pressure	yes	.09	.01	13.64	1	<.001
	no	.16	.03			
sns*rspeed	figure 28(a)			7.59	1	.006
sns*trust	figure 28(b)			13.97	1	.<001

Moreover, the fixation data shows that there is a significant interaction of SNS*trust and SNS*reading speed as shown in figure 28(a) and 28(b). For good readers, the percentage of fixation on result scene stayed the same with and without SNS, but average readers had a significantly lower number of fixations on the results under the SNS condition. Participants with low trust had a significantly lower number of fixations on result under the SNS condition, while the number of fixations increased for participants with high trust. This does not match my intuition that high trust participants would rely more on SNS and would have less fixations on results



(a) SNS and reading speed



(b) SNS and trust

Figure 28: Analysis of percentage of fixations on *results* AOI

Fixation duration on search results: Fixation duration is an eye tracking measure of information processing difficulty. As mentioned earlier, longer fixations represent more information processing that can be due to higher density of information or more difficult information [49], [35]. I analyzed average fixation duration on the results AOI for all four sessions. The result of analysis is shown in table 19. There is a significant effect or order, and topics. The third session has significantly longer fixations and average fixation duration for topic 4 is significantly higher than topic 1, 2, and 3. Significant interaction of SNS and time pressure shows that under time pressure with SNS, the average fixation duration decreased significantly. This suggests that SNS decreased the difficulty of finding information under time pressure. Additionally, SNS decreased fixation duration for low trust group, average readers, and female participants.

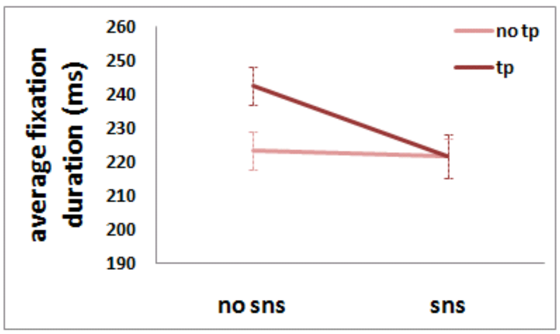
Table 19: Analysis of fixation duration

		μ	SE	Wald χ^2	df	p-value
order	1	216.83	8.28	62.317	3	<.001
	2	220.73	7.07			
	3	252.88	7.3			
	4	220.25	6.01			
topic	1	223.85	6.4	29.035	3	<.001
	2	221.78	4.71			
	3	221.54	8.36			
	4	242.37	6.30			
sns*tp		figure 29(a)		8.255	1	.004
sns*trust		figure 29(b)		8.255	1	.004
sns*gender		figure 29(c)		5.898	1	.015
sns*reading speed		figure 29(d)		16.710	1	.000

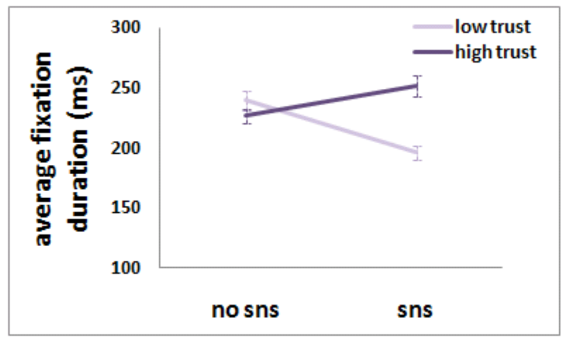
4.2.4 Summary of findings

The results support that participants follow SNS more under time pressure. They access SNS-documents significantly more and have a significantly higher number of fixations on social navigation icons under time pressure. Contrary to my expectation, there is no strong evidence that female participants follow social navigation more. Under no time pressure, female participants had significantly more fixations on highlighted text; however, it drops significantly under time pressure and male participants have significantly more fixations on highlighted text under time pressure. Additionally, male participants collected significantly more notes from SNS-documents.

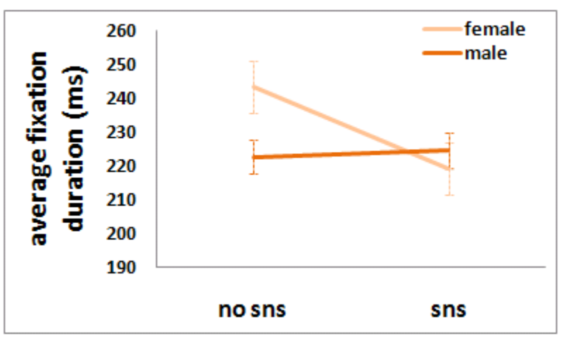
Readers with good speed seem to follow social navigation support the most. They collected significantly more notes from SNS-documents, accessed significantly more documents from the maps, and had significantly higher number of fixations on the maps. Time pressure has a different effect on participants with different reading abilities. While good readers made use of social navigation cues independent of time constraint, excellent readers accessed SNS-



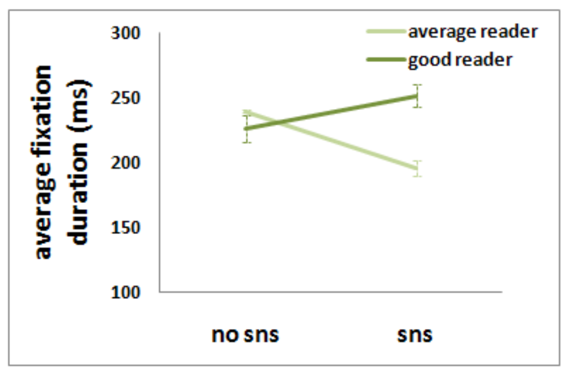
(a) SNS and time pressure



(b) SNS and trust



(c) SNS and gender



(d) SNS and reading speed

Figure 29: Analysis of fixation duration on *results* AOI

documents significantly more under time pressure, and average readers had a significantly higher number of fixations on highlighted passages under time pressure.

Interpersonal trust plays a role in the following of SNS. While both high and low trust groups had a similar number of fixations on maps, participants with high interpersonal trust accessed significantly more documents from the maps. This result matches the utility of the maps since the decision to access a document from the maps is made primarily based on activities of others. On the other hand, when participants were limited by time pressure, the number of fixations on highlighted text for the low trust group increased significantly.

Additionally, the results show evidence that SNS decreases participants' efforts in some cases. good readers and the high trust group, who followed SNS the most, entered significantly fewer queries under the SNS condition. Participants used shorter queries under time pressure when SNS was provided. Additionally, the high trust group and the female participants used shorter queries with SNS.

Average fixation duration on results decreased significantly under time pressure when SNS was offered. Female participants and average readers had lower average fixation duration on results with SNS. However, average fixation duration is influenced by the trust factor in reverse order of my expectations. It increased for the high trust group and decreased for the low trust group under the SNS condition. This can be the result of the low trust group possibly ignoring social navigation cues more; as a result, under the SNS condition high trust group had more information to process and therefore had longer fixation duration. There was no significant effect of SNS on collecting note rate and time spent searching.

4.3 EFFECT OF SNS ON TASK PERFORMANCE

The previous section presented a significant effect of SNS on different aspects of users' search behaviors. In this section I present whether or not the difference in behavior resulted in different performance. The ideal approach in measuring the performance would have been asking at several judges to assess how relevant are the passages the participants had annotated to the task and how complete they had performed the task. However, the approach is labor and time intense and was not an option in this study. In the current study the performance was measured by comparing the notes against "group truth" and judging whether a note was relevant to the task. Using the "ground truth," I measured the task performance at the level of documents and notes.

4.3.1 Definitions

- **Relevant document**: A document is defined as relevant if it includes at least one note matching the ground truth data.
- **Relevant note**: A note is defined as relevant if it overlaps with a note from the ground truth with some specific threshold.

4.3.2 Measurements

- **Document level performace**: document level performance is measured by the classical information retrieval F-measure, which takes into account both precision and recall. Recall is specifically important for this task since the participants were explicitly asked to find as much relevant information as time allows. It is defined as:

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

where

$$Precision = \frac{number\ of\ relevant\ visited\ documents}{total\ number\ of\ visited\ docuemnts}$$

and

$$Recall = \frac{number\ of\ relevant\ visited\ documents}{total\ number\ of\ relevant\ documents}$$

- **Note level performance**: Ground truth includes annotations by two annotators; therefore, the overlap can happen with the union or intersection of ground truth notes. Considering the strength of overlap and three different thresholds of overlap, I defined six different relevance values for each note collected by each user. I considered 50%, 75%, and 90% as the thresholds for defining the relevance. After assigning 1 or 0 to each note for each session, I computed the percentage of relevant notes by dividing the number

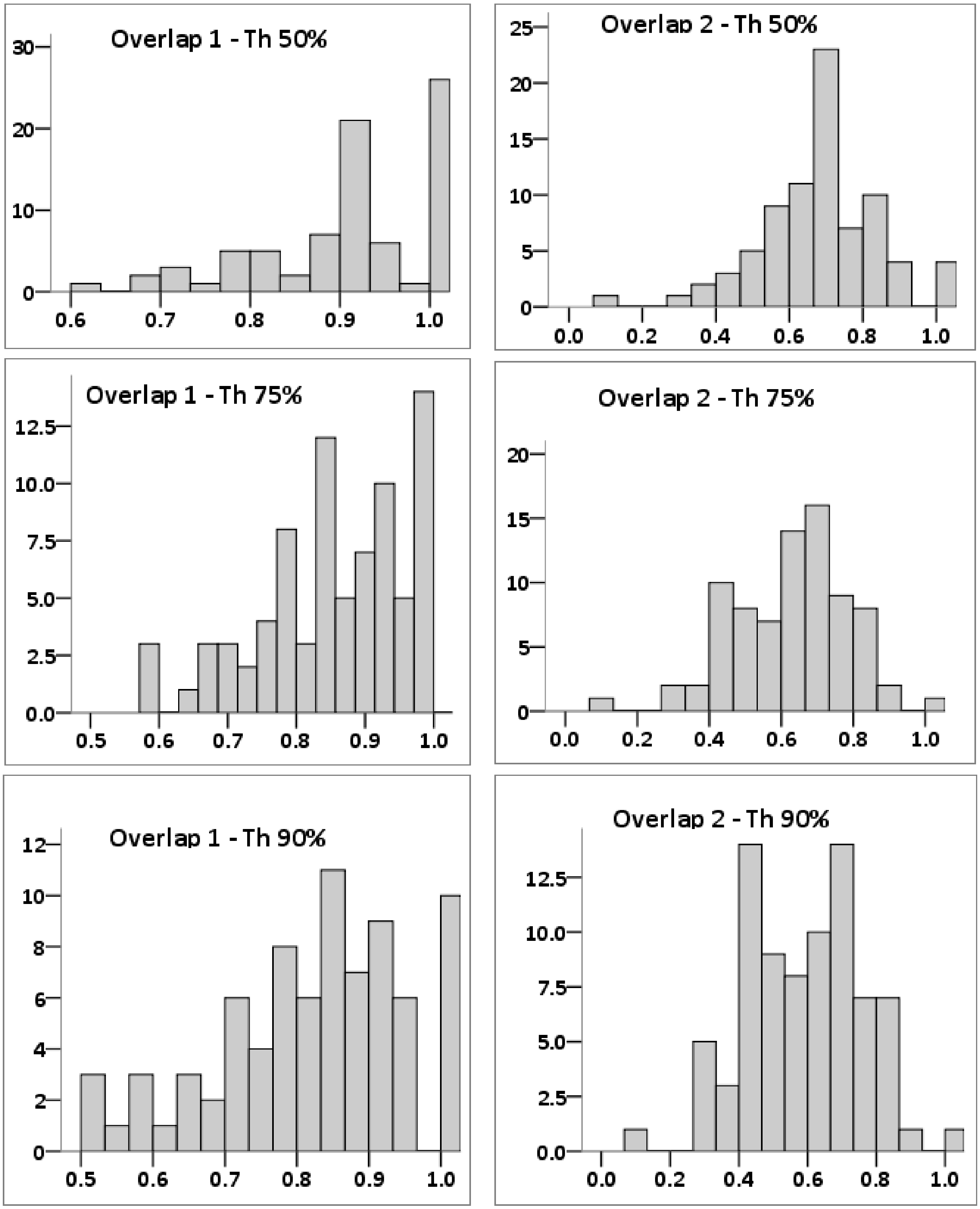


Figure 30: Distribution of different relevance values for all collected note

of relevant notes (notes with relevancy value of 1) to the total number of notes. The distribution of different relevance is shown in figure 30. The left side of the figure shows the distribution of relevant notes when overlap with at least one ground truth annotator is considered (union) and the right side is the relevance data when the overlap with both annotators are taken into account (intersection). The X axis shows the percentage of relevant notes and Y axis shows the number of sessions. The most liberal choice is overlap 1 with 50% threshold. As a result there is a bias toward having two many relevant notes. The most strict choice is overlap 2 with 90% threshold which has a bias toward having very few relevant notes. Overlap 1 with 90% threshold and overlap 2 with 75% threshold have the most uniform distributions. Those two cases were used for analysis.

4.3.3 Quality of Social Navigation Cues

The result of the previous section showed that participants' information seeking behavior was influenced by social navigation cues. Therefore, before looking into performance, it is important to assess the quality of social navigation cures. I assessed the quality of social navigation cues similar to assessment of performance. At the document level, the result shows that for all four topics all documents augmented with social navigation cues had at least one relevant note from ground truth collection.

Table 20 shows the percentage of relevant passages for each topic given 90% threshold with at least one annotator, and 75% threshold with both annotators criteria. The relevancy is calculated by dividing the number of relevant notes responding the questions in the task for the current participants by total number of notes they could have seen. The data shows that the note relevancy with both criteria is low for all four topics. That means, while social navigation cues could guide them to relevant documents, the highlighted passages were not relevant to their task.

4.3.4 Document based Task Performance

As mentioned earlier, F-measure was used to assess document-based task performance. The result of the analysis is shown in table 21. There is a significant effect of SNS. With SNS

Table 20: Quality of social navigation cues

topic	TH 90	TH 75
40001	34%	22%
40038	12%	10%
41012	15%	15%
41019	19%	19%

participants achieved a significantly higher F-score which means they found significantly more relevant documents. There is a significant effect of reading speed which shows that excellent readers achieved a significantly higher F-score while good and average readers performed the same.

Table 21: Analysis of F-measure

		μ	SE	Wald χ^2	df	p-value
sns	yes	.113	.003	5.79	1	.016
	no	.106	.002			
reading speed	average	.095	.005	15.587	2	<.001
	good	.098	.004			
	excellent	.145	.013			
time pressure*topic		figure 31		13.55	3	.004

Additionally, there is a significant interaction of time pressure and topic as shown in figure 31. Naturally under no time pressure, participants achieved a higher F-score because they had double the time. The result shows that under time pressure participants achieve a similar F-score on all topics, while under no time pressure, they achieved a significantly higher F-score for topic 2 and 4.

4.3.5 Note based Task Performance - 90% Overlap with at Least One Annotator

The result of the analysis of note relevance considering 90% overlap with at least one of the ground truth annotators is shown in table 22. There is a significant effect of topic, and

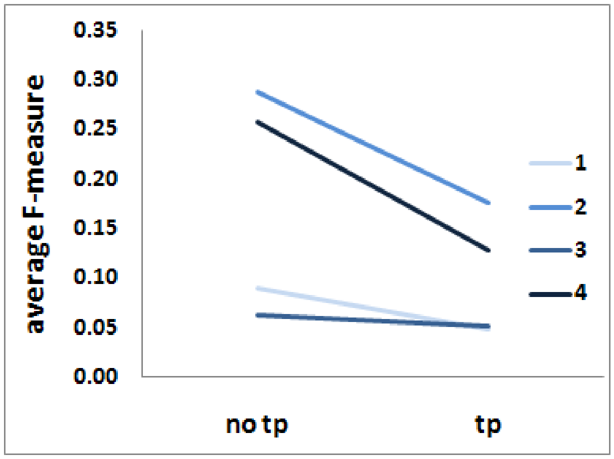


Figure 31: Effect of interaction of time pressure and topic on participants' document-level performance

Table 22: Analysis of note relevance: overlap1-th90%

		μ	SE	Wald χ^2	df	p-value
topic	1	.85	.02	60.45	3	<.001
	2	.90	.02			
	3	.74	.02			
	4	.78	.02			
trust	high	.85	.02	4.16	1	.041
	low	.78	.02			
sns * time pressure				3.473	1	.062

trust. Participants found significantly higher number of relevant notes for topic 1 and 2 compared with topic 3 and 4. Participants with high trust collected a significantly higher number of relevant notes. There is no significant effect of SNS. The interaction of SNS and time pressure is marginally significant, which suggests a pattern of collecting more relevant notes under time pressure when SNS was provided.

4.3.6 Note based Task Performance - 75% Overlap with Both Annotators

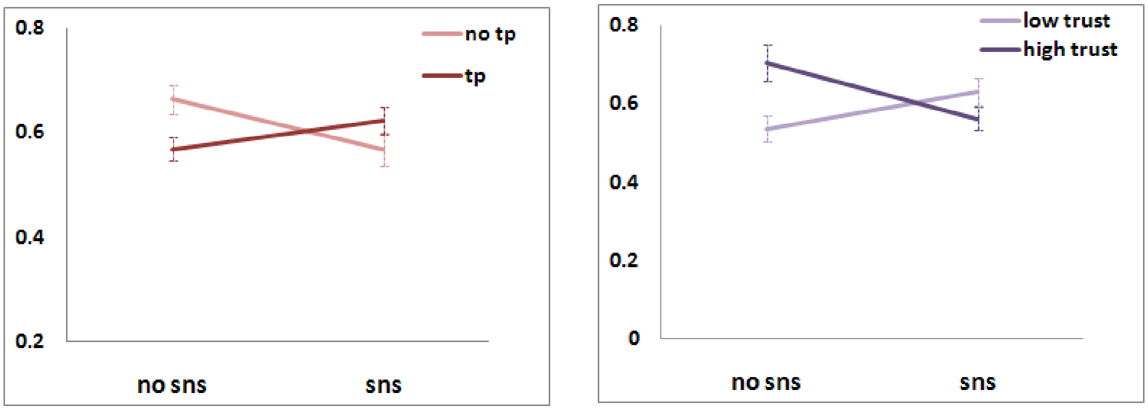
The result of the analysis of the note relevance considering overlap with both ground truth annotators and 75% threshold is shown in table 23. Using more strict criteria, there is a significant interaction of sns*time pressure and a significant interaction of sns*trust. Figure 32(a) presents the interaction of time pressure and SNS. The result shows that when SNS was provided, participants collected slightly more relevant notes under time pressure and significantly less relevant notes under no time pressure. The interaction of trust and SNS is shown in figure 32(b). It was expected that participants with high trust collect more relevant notes with SNS; however, the result shows participants with high trust collected less relevant notes under the SNS condition.

Table 23: Analysis of note relevance: overlap2-th75%

	Wald χ^2	df	p-value
topic	34.880	3	.000
sns * time pressure	6.114	1	.013
sns * trust	13.964	1	.000

4.3.7 Summary of Findings

The analysis of document level performance of participants shows that SNS helped the participants to find more relevant documents. The effect of SNS at the note level was less evident. No effect of SNS was observed when less strict criteria of note relevancy was applied. Applying the more strict criteria to require a relevant note to overlap with both ground truth annotators revealed some effect of SNS. While SNS helped the participant



(a) SNS and time pressure (b) SNS and trust

Figure 32: Effect of SNS on note relevance considering overlap2-th75% criteria

boost their performance under time pressure, it harmed their performance under no time pressure. Additionally, participants with high trust which were more likely to follow SNS, performed worse in terms of relevancy of notes they had collected.

The result matches the design of SNS in the experiment. SNS was designed to guide the participants to relevant documents; however, at the document level, the highlighted text did not necessarily help the participants collect more relevant notes since the passages to related questions and not exactly the same questions were highlighted.

4.4 USERS' SUBJECTIVE OPINIONS

In addition to log analysis and eye tracking, I collected participants' opinions on task difficulty and SNS features of the application. I was specifically interested to study whether there is any correlation of personal factors and experimental conditions on their opinion. As explained in the procedure of the study, they responded to a questionnaire at the end of each session. There were three versions of questionnaires depending on the condition of the session (Appendix D):

- Five-question version after sessions with no SNS: This version included general questions about task difficulty using the application and whether they had enough time to perform the task.
- Six-question version after sessions with no SNS following an SNS session: This version included the general questions plus one extra question asking whether participants found it easier to find information with SNS.
- 13-question version after session with SNS: This included the five general questions plus eight questions asking participants' opinions about different social navigation features.

The possible responses to all questions ranged from 1 (not at all) to 5 (extremely). Therefore the higher the number, the more positive the response.

4.4.1 Evaluating Task Difficulty

Table 24 shows the result of analysis of sum of responses to the first four general questions. The average is 14 (SD=3.478). There is a significant effect of order. The responses for the first session are significantly more positive than other sessions. There is also a significant effect of topic. The responses for topic 2 are significantly more positive than other topics and responses for topic 3 are significantly more negative.

Table 24: Analysis of participants' opinion about task difficulty

		μ	SE	Wald χ^2	df	p-value
order	1	16.15	.693	27.73	3	<.001
	2	12.99	.845			
	3	14.75	.831			
	4	14.18	.705)			
topic	1	15.08	.76	58.58	3	<.001
	2	16.88	.86			
	3	11.60	.70			
	4	14.86	.78			
sns*reading speed				8.90	2	.012

The significant interaction of sns and reading speed is shown in figure 33. Good readers have significantly more positive feedback under the SNS condition while excellent readers have significantly more negative feedback under the SNS condition.

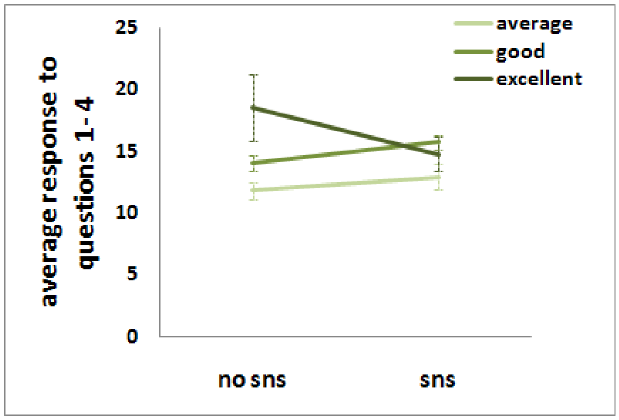


Figure 33: Effect of SNS on participants' opinion about task difficulty with different reading abilities

4.4.2 Evaluating Time Constraint

Table 25: Analysis of participants' opinion about time constraint

		μ	SE	Wald χ^2	df	p-value
topic	1	3.45	.33	38.43	3	<.001
	2	4.23	.32			
	3	2.26	.23			
	4	3.27	.24			
time pressure	yes	2.63	.17	45.74	1	<.000
	no	3.94	.28			
reading speed	average	2.55	.20	11.87	2	.003
	good	3.45	.16			
	excellent	3.80	.59			
sns*time pressure				4.263	1	.039

Question 5 asked the participants' opinions about time pressure. The result is shown in table 25. There is a significant effect of time pressure, which is the ensuring result that participants felt under time constraint in the shorter sessions. There is a significant effect of reading speed which shows that average readers felt significantly more time constrained. There is an effect of topic. Participants felt the time constraint significantly more for the third topic and significantly less for the second topic. There is also a significant interaction of SNS and time pressure as shown in figure 34. Participants found the SNS sessions similar in terms of time constraint.

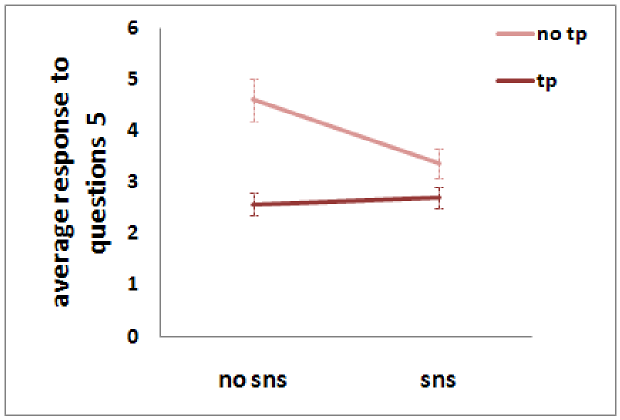


Figure 34: Effect of SNS on participants' opinions about time constraint

4.4.3 Evaluation Social Navigation Features

Question 8 asked participants whether they found it useful to know what documents were selected by other users. The average response is 3.40 (SD=1.32). The result of analysis is shown in table 26. There is a significant effect of order, topic, reading speed, and trust on participants' response to the question. Participants had significantly more positive response for first and last sessions . They had more negative responses for topic 3 compared with topic 1, 2 and 4. Average readers had the most positive responses and good reader significantly more positive than excellent readers. Surprisingly, participants with low trust had significantly more positive responses.

Question 9 asked the participants whether they found it useful to know what documents were annotated by others. The average response is 3.40 (SD=1.34). The result of analysis is shown in table 27. There is a significant effect of topic, reading speed, and trust on participants' responses to the question. Similar to responses to question 8, participants

Table 26: Analysis of participants' opinions about social navigation features: Q8

		μ	SE	Wald χ^2	df	p-value
order	1	3.01	.33	9.80	1	<.001
	2	2.54	.33			
	3	2.12	32			
	4	3.05	.40			
topic	1	2.97	.34	64.95	3	<.001
	2	3.27	.42			
	3	1.77	.22			
	4	2.86	.38			
reading speed	average	4.79	.46	16.49	2	<.001
	good	2.99	.24			
	excellent	1.30	.44			
trust	high	1.99	.299	4.73	1	.030
	low	3.52	.452			

Table 27: Analysis of participants' opinion about social navigation features: Q9

		μ	SE	Wald χ^2	df	p-value
topic	1	2.65	.29	30.57	3	<.001
	2	3.20	.44			
	3	2.04	.25			
	4	2.99	.43			
reading speed	average	4.87	.51	17.31	2	<.001
	good	2.84	24			
	excellent	1.39	47			
trust	high	2.19	.34	10.704	1	.001
	low	3.29	.43			

had significantly more negative response for topic 3 comparing with topic 1, 2 and 4. Average readers had significantly more positive responses and good reader significantly more positive than excellent readers and participants with low trust had significantly more positive responses.

Question 10 asked participants whether they found it useful to view highlighted passages by other users. The average response is 3.25 (SD=1.35). The result of the analysis is shown in table 28. There is a significant effect of topic, reading speed, and gender on participants' responses to the question. Participants had significantly more negative response for topic 3 compared with topic 1, 2 and 4. Average readers had significantly more positive responses and good reader significantly more positive than excellent readers. Male participants had significantly more positive responses.

Table 28: Analysis of participants' opinion about social navigation features: Q10

		μ	SE	Wald χ^2	df	p-value
topic	1	2.71	.31	19.39	3	<.001
	2	2.82	.40			
	3	1.99	.20			
	4	2.64	.32			
gender	female	2.09	.27	6.32	1	.012
	male	3.04	.36			
reading speed	average	4.73	.48	22.62	2	<.001
	good	2.70	.26			
	excellent	1.25	.39			

Question 13 asked participants' opinions about social maps. The average response is 3.46 (SD=1.2). The result of the analysis is shown in table 29.

There is a significant effect of topic, gender, reading speed, and trust on participants' responses to the question. Participants had significantly more positive response for topic 1 compared with topic 2, 3 and 4 and more negative responses for topic 2 compared with topic 1 and 4. Average readers had significantly more positive responses and good reader significantly more positive than excellent readers . Participants with high trust had significantly

Table 29: Analysis of participants' opinions about social navigation features: Q13

		μ	SE	Wald χ^2	df	p-value
topic	1	3.14	.13	84.983	3	<.001
	2	2.54	.17			
	3	1.99	.09			
	4	2.13	.16			
gender	female	2.69	.11	9.203	1	.002
	male	2.16	.11			
reading speed	average	4.30	.26	411.793	2	<.001
	good	3.77	.17			
	excellent	.86	.08			
trust	low	2.01	.16	9.862	1	.002
	high	2.89	.13			

more positive responses. Male participants had significantly more positive responses.

4.4.4 Summary of the Findings

The participants' opinions show that short sessions indeed left them with little time to complete the task. Specifically, participants with the lowest reading abilities expressed having not enough time to find the information for which they were asked. This result confirms the design of the study in terms of time allocation and also the reading speed criteria used to classify participants based on their reading abilities. On the other hand, the significance of the difference of opinion as a result of time pressure disappears when SNS was provided.

Overall the difficulty of the task was rated medium by the participants. Good readers who showed the sign of following SNS the most had found the task less difficult when SNS was provided. Overall attitude toward SNS features was positive and, specifically, participants with low or good reading speed appreciated all forms of SNS: augmenting icons, social maps, and highlighted text. Surprisingly, participants with lower interpersonal trust appreciated SNS more. Gender effect appeared only when the participants where asked about the in-text highlighted passages, and social maps, and in both cases, male participants appreciated the features more.

5.0 DISCUSSION OF THE RESULTS

In this chapter, I revisit the main hypotheses of the study and discuss the supporting evidence for each hypothesis. Figure 35 shows the overall result of analysis. It presents the significant effect of different factors and their interactions on all measures. Cells with dark green background present significant differences at α level of .01, cells with light green shows significant differences at α level of .05, and very light green is a representative of marginally significant differences (α=.1). Eye tracking measures are highlighted in orange background.

et measures | α=.01 | α=.05 | α=.10

		Main Effects							Interactions						
		topic	order	tp	sns	gender	trust	rspeed	tp			sns			tp
									gender	trust	rspeed	gender	trust	rspeed	sns
Following SNS	SNS-documents access	3<1,2,4		tp>no tp	na										
	SNS-documents collect note				na	m>f		good>average, excellent							
	Maps access rate			no tp>tp	na		high>low	good>average, excellent							
	normalized fc on maps			no tp>tp	na	m>f		good>average, excellent							
	normalized fc on icons			tp>no-tp	na	f>m	low>high								
	normalized fc on htxt	2>1,3,4	2,3>1,4		na										
Effort	number of queries	1,3,4>2		no tp>tp											
	query length	1,2,3>4													
	Collect note rate			no tp>tp				excellent>good, average							
	Time spent searching	1,3,4>2	1,2>34	no tp>tp				good>average, excellent							
	normalized fc on results	1,3,4>2		no tp>tp											
	Fixation duration on results	4>1,2,3	3>1,2,3												
performance	document level performance				sns>no sns			excellent>good>average							
	note level performance (1)	2>1,3,4					high>low								
	note level performance (2)	2>1,3,4													
participants' opinion	Task difficulty	1,2,4>3	1>2,3,4												
	time constraint	2>1,3,4		no tp>tp				excellent>good>average							
	Social navigation features	1,2,4>3	1,2,4>3		na		low>high	average>good>excellent							

Figure 35: Summary of analysis

5.1 EFFECT OF TIME PRESSURE

Hypothesis 1 - Under time pressure, participants utilize social navigation support more as an extra support to cope with time constraint. Suggested by information foraging theory, participants are likely to use social navigation cues as information scent to detect the importance and relevance of an information item. The need to use the social information scent is intensified by time pressure. Therefore, participants are expected to pay more attention to social navigation cues (icons and maps) and follow the cues more.

Eye movements measures provided information about how much social navigation cues were noticed under different conditions. Participants had significantly higher number of fixations on social navigation icons under time pressure. Consistently, the analysis of their click-stream shows that participants clicked significantly more on documents with social navigation icons.

On the other hand, there was no significant effect of time on how much they looked at social maps. The number of fixations on the maps and access rate from the maps were non-significantly higher under no time pressure. Social maps provide an alternative way of navigation which is highly dependent on social navigation support. Frequently, users perform information search by issuing a query and reviewing a list of results including title, and abstract for each item. Unfamiliarity with information access method through the maps might have played a role in less attention under time pressure. It is feasible that under time pressure, the participants tried to use an information seeking strategy which is more familiar for them. The order number of fixations on social maps under no time pressure provides supporting evidence. Majority of fixations on social maps happened towards the end of the session;i.e., participants started paying attention to the maps after exploring and possibly exhausting exploration of the results specially high ranked results.

Following social navigation support under time pressure affected the effort and performance of the participants. They entered significantly shorter queries which indicates that shorter queries were sufficient to get the desired information. They utilized social navigation support to cope with time constraint by avoiding elaboration of search queries. Moreover, social navigation support guided the participants to relevant documents which resulted in accessing significantly more relevant documents.

An interesting observation is perception of time constraint by the participants under different conditions. While the result confirms that participants did not have enough time under time pressure to perform the task, with social navigation support even "no time pressured" sessions were not perceived as long enough. This suggests that social navigation support might require extra processing. Moreover, with social navigation support they are able to see all the documents that were accessed by prior users. Therefore, participants are likely to think that they had not accessed and reviewed all relevant information. This can result into perception of not enough time.

Overall, the results are in some degrees supportive of the first hypothesis that social navigation support is used more under time pressure; however, it is important to take into account the way the support is provided. Unfamiliarity with the presentation and implementation approach can affect the utility of social navigation support.

5.2 EFFECT OF INTERPERSONAL TRUST

Hypothesis II - Participants with high interpersonal trust are more likely to use social navigation support. Social navigation support is based on activity of other users; therefore, it is expected that users with higher interpersonal trust rely more on social navigation cues and follow them more.

Analysis of fixations data shows that social navigation cues were not ignored by low trust group and in fact low trust group had higher number of fixations on social navigation icons. However, access to social maps was significantly higher by high trust group. This result shows that while both group paid attention to the social navigation icons and maps, low trust group did not rely on and make use of them as much as high trust group. Additionally, eye movement data shows that high trust group have larger number of fixations with longer duration on results. This finding is consistent with perception of time under social navigation support which suggest that social navigation support may require extra processing causing higher number of fixations with longer duration.

The analysis of the task performance shows that participants with high interpersonal trust collected lower percentage of relevant notes with social navigation support. The drop of performance might be a result of high trust on social navigation cues. Highlighted passages were at highest 34% relevant to the task. As a result, if note collection behavior of participants was highly influenced by highlighted passages, they could have collected irrelevant notes by following traces of others.

The result shows that trust plays an important role on the effect of social navigation support of users' information seeking behaviors. This is specifically important for designer of applications with social navigation support to ensure high reliability of social navigation cues. It is important to use all possible measures to avoid guiding the users to the wrong direction. Presenting the uncertainty associated with social cues when high reliability cannot be achieved can serve as a solution.

5.3 EFFECT OF GENDER

Hypothesis III - Female participants are more likely to follow social navigation support. Research on gender differences and navigation in real world and hypermedia shows that females are more likely to feel disoriented and be in need of navigation support.

The result of current study does not provide any evidence that female participants utilize social navigation support more than male participants. Eye movement analysis shows that there is no significant difference in terms of number of fixations on social navigation icons, social maps, or highlighted passages. Also there is no difference on the number of documents accessed from the maps. In terms of accessing documents augmented with social navigation icons, male participants accessed significantly more of those documents. This finding is dependent on the type of the task. Informational search tasks in general and the particular task of the current study required minimum navigation and mostly the participants had to explore and read the text to find the relevant information. Therefore, it is not very likely that females participant felt disoriented. On the other hand, male participants might have been attracted to social navigation support for reasons other than feeling of disorientation.

A study of social navigation support for a navigational task can help to achieve better understanding of the effect of gender on following social navigation support.

Prior studies of gender differences in web search shows that females tend to use longer queries. The result of current study confirms that. With no social navigation support, females entered significantly longer queries; however, the difference was eliminated with social navigation support. This suggests that social navigation support may have increased the confidence and trust of female participants in the system saving them longer elaboration of queries.

5.4 EYE MOVEMENT ANALYSIS

Employing eye tracking and analysis of eye movement data provided additional hints to understand users' information seeking behaviors. In particular, it provided information about how much social navigation cues are noticed by different participants and under different conditions and whether their attention to social navigation cues correlated with their actions.

Analysis of eye movement data contributed to assessment of effect of time pressure, interpersonal trust, and gender on following of social navigation support. It showed that under time pressure, users pay more attention to social navigation cues embedded with the results while they do not pay attention to social navigation support offered as a new navigation method. Users mainly attended to social maps were they had additional time. Eye movement data showed that less usage of social navigation support by low trust group is not an indication of not spotting those cues. In contrary, low trust group attend to those cues as much as high trust group and not making extensive use of them is a decision make with awareness. Similarly, unexpected low usage of social navigation support by female participants is not due to not noticing the icons or maps. The data shows that number of fixations on maps and icons are similar for both females and males.

Additionally, analysis of fixation duration provided evidence for additional processing required for social navigation support. Participants' opinion about time suggests that they needed more time to process social navigation cues and longer fixation duration on results with social navigation support confirms that. Longer fixation duration is frequently associated with higher information processing. This suggests that participants had to encode and process social navigation information and the relation between each result item and icons associated to them which caused longer fixations.

6.0 CONCLUSION

This dissertation presented a multifaceted study of SNS in a controlled experiment designed for factual information seeking tasks. Eye movements data helped to gain more insight into the effect of SNS on users' search behaviors. The result of the study shows that social navigation cues affect users' search behaviors and users pay attention to those cues and follow them for finding information. However, a different kind of navigation support is utilized by different users. Navigation cues embedded with traditional search strategies were more likely to be employed by female and low trust participants while additional support provided through social maps was more likely to be used by male and high trust participants. SNS was successful in guiding participants to the relevant documents, assisting them to boost their document level performance.

Moreover, the result of the study provided evidence that SNS interacts with time pressure and under time pressure users are more likely to follow SNS. An interesting finding of the study is that the "snowball effect" often associated as limitation of SNS can be controlled by providing content based information to allows users to make their own informed judgments. In this study, the participants were not attracted to highlighted text by prior users blindly and they read the non highlighted part of the text as much as the highlighted part. Irrelevant notes did not harm their performance as long as they had enough time to read the text. However, this was affected by time pressure. Under time pressure participants paid more attention to highlighted passages and possibly collected some notes from irrelevant passages which harmed their performance. The effect was stronger on participants who relied more on SNS and followed SNS more.

Reading abilities of the participants had an important effect on how much they followed SNS. Participants with a good reading speed were more likely to follow SNS. Interpersonal trust also played a significant role on search behavior and users' opinions of the system. The results of the study suggest to take into account these factors while designing SNS. Gender played a significant role on limited variables. The effect was less apparent than expected and found in general navigation and Web navigation literature.

6.1 LIMITATIONS OF THE STUDY

The experiment was conducted with a limited number of participants and under specific circumstances designed by the study. Generalizations of this work should be done with careful considerations. Specifically eye tracking data analysis included a lower number of participants. A larger number of participants could help to strengthen the findings of this study. A bigger sample size could also allow to take into account other personal characteristics that may affect the findings such as sociability of participants.

The collaborative nature of the task in this study was hidden from the participants which helped to study indirect and unintentional SNS. However, as mentioned in the literature review, this is only one way of providing SNS. It would be interesting to study the effect of different types of SNS in a more perceivable collaborative task.

Interpersonal trust was measured with a questionnaire which was designed for measuring trust in both acquainted and stranger peers. In the current study, the participants had no information about other users who were the source of SNS. A questionnaire focused on measuring interpersonal trust in strangers could have been more appropriate for the task. The decision to use the current questionnaire was made to avoid validating a novel questionnaire and make use of what has precedence in the literature. On the other hand, SNS is often used in closed communities of users who are acquainted; it is interesting to study the effect of SNS while people have social contacts with other users leaving traces.

SNS in the current study was offered in two forms: Enhancing search results and providing social maps that offers a different type of information navigation strategy. The design does also not allow to discriminate the effect of navigation strategy and navigation support which might have affected some findings of the study.

APPENDIX A

DEMOGRAPHIC QUESTIONNAIRE

1. Age

☐ <20 ☐ 20 – 25 ☐ 25 – 30 ☐ 30 – 35 ☐ > 35

2. Gender: ☐ Female ☐ Male

3. Mother tongue:

4. University: ☐ University of Pittsburgh ☐ Carnegie Mellon

5. Degree:

☐ Undergraduate ☐ Masters ☐ PhD

6. Major:

7. SAT/GRE/CAT score:

Verbal:________________ Overall:_________________

8. Reading speed

☐ Slow ☐ Average ☐ Good ☐ Excellent

9. Have you taken any information retrieval course?

If yes, what is the course title?

10. How often do you use Internet search such as Google?

☐ Several times a day ☐ Few times a day
☐ Several times a week ☐ Few times a week

11. Do you use any search engines other than Google?

If yes, which ones?

12. Are you a member of any social networking sites such as Facebook?

If yes, which ones? And how often do you login to the sites?

☐ More than once a day ☐ Once a day
☐ Few times a week ☐ Once a week
☐ Less than once a week

13. Do you use any social bookmarking site such as Delicious, Digg, or Flickr? If yes, which ones?

Figure 36: Demographic Questionnaire

APPENDIX B

INTERPERSONAL TRUST QUESTIONNAIRE

I tend to be cynical and skeptical of others' intentions

☐ Strongly Agree
☐ Agree
☐ No Opinion
☐ Disagree
☐ Strongly Disagree

I believe that most people will take advantage of you if you let them

☐ Strongly Agree
☐ Agree
☐ No Opinion
☐ Disagree
☐ Strongly Disagree

If I got into difficulties at work I know my colleagues would try and help me out

☐ Strongly Agree
☐ Agree
☐ No Opinion
☐ Disagree
☐ Strongly Disagree

I can trust the people I work with to lend me a hand if I needed it

☐ Strongly Agree
☐ Agree
☐ No Opinion
☐ Disagree
☐ Strongly Disagree

Most of my peers can be relied upon to do as they say they will do

☐ Strongly Agree
☐ Agree
☐ No Opinion
☐ Disagree
☐ Strongly Disagree

Figure 37: Interpersonal Trust Questionnaire

APPENDIX C

TASKS DESCRIPTION AND BACKGROUND KNOWLEDGE QUESTIONNAIRE

C.1 TOPIC - EARTHQUAKE HIT INDIA'S GUJARAT STATE

A huge earthquake hit India's Gujarat state, January 26, 2001. The task is to find the information about what rescue and relief actions have be taken on January 27, 2001, the day after the earthquake.

C.1.1 Background Knowledge

C.1.2 Task Description

A huge earthquake hit India's Gujarat state, January 26, 2001. The task is to find the information about what rescue and relief actions have be taken on January 27, 2001, the second day after the earthquake. From the articles, find snippets of relevant text to the following questions:

1. **Where did the earthquake happen?**
2. **What was the number of injure?**
3. **How many troops were sent in?**
4. **How many tents and other materials were needed?**
5. **How many relief materials were sent in?**

1. What is your level of knowledge about this story?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

2. Do you recall reading a newspaper article (either in print or online), or listened to or watched a news clip about this story?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

3. What is your level of knowledge about India?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

4. What is your level of knowledge about international organizations such as United Nations?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

Figure 38: Topic 1 - Background Knowledge Questionnaire

And here is the list of questions you colleagues have worked on:

1. What was the degree of the earthquake?
2. When did the earthquake happen?
3. What was the number of death?
4. How much money lost?
5. How about the lack of medical facilities?

C.2 TOPIC - GALAPAGOS OIL SPILL

An Ecuadorian oil tanker spilled fuel near Galapagos Islands and the fuel threatened local beaches, wildlife and impacted fishing. The task is to suggest on where will the current leaks of the oil go, what actions should to be taken after the fuel-leaking event caused by the oil tanker Jessica near the Galapagos Islands, as well as possible support that can be provided by US. You might also be interested in possible law cases towards the responsible persons.

Note that tanker ran around on Jan. 16, but started to leak on Jan. 19. The earliest report we have is on Jan. 20.

C.2.1 Background Knowledge

1. What is your level of knowledge about this story?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

2. Do you recall reading a newspaper article (either in print or online), or listened to or watched a news clip about this story?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

3. What is your level of knowledge about Ecuador?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

4. What is your level of knowledge about wildlife and fishing?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

Figure 39: Topic 2 - Background Knowledge Questionnaire

C.2.2 Task Description

An Ecuadorian oil tanker spilled fuel near Galapagos Islands and the fuel threatened local beaches, wildlife and impacted fishing. The task is to suggest on where will the current leaks of the oil go, what actions should to be taken after the fuel-leaking event caused by the oil tanker Jessica near the Galapagos Islands, as well as possible support that can be provided by US. You might also be interested in possible law cases towards the responsible persons. Note that tanker ran around on Jan. 16, but started to leak on Jan. 19. The earliest report we have is on Jan. 20. From the articles, find snippets of relevant text to the following questions:

1. How much gallons fuel emitted?

2. **What is the cause of the leaking event?**
3. **Who is responsible for leaking?**
4. **What animals are affected by the leaked fuel?**
5. **Who put efforts in relieving the pollution?**

And here is the list of questions you colleagues have worked on:

1. What is the date of fuel leaking?
2. Where is the location of fuel leaking?
3. What type of vehicle is in the event?
4. Any support from other countries?
5. What is the impact on the local people, environment, animals and plan?

C.3 TOPIC - TROUBLE IN THE IVORY COAST

The task is to find information on the elections in the Ivory Coast in 2000/2001 and focus on the potential for civil war.

C.3.1 Background Knowledge

C.3.2 Task Description

The task is to find information on the elections in the Ivory Coast in 2000/2001 and focus on the potential for civil war.

From the articles, find snippets of relevant text to the following questions:

1. **What crimes has Guei been accused or convicted of?**
2. **What are the roles of the UN and France in the elections & following riots?**
3. **What were the demonstrations following election?**
4. **What was the outcome of election outcome?**
5. **What was the ethnic, political and religious tensions?**

1. What is your level of knowledge about this story?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

2. Do you recall reading a newspaper article (either in print or online), or listened to or watched a news clip about this story?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

3. What is your level of knowledge about Ivory Coast?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

4. What is your level of knowledge about civil wars in Africa?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

Figure 40: Topic 3 - Background Knowledge Questionnaire

And here is the list of questions you colleagues have worked on:

1. Why did Robert Guei flee?
2. Where is Robert Guei?
3. How many people were injured in riots?
4. How many people were killed in riots?
5. What was the cause of violence?

C.4 TOPIC - ILIESCU WINS ROMANIAN ELECTIONS

Iliescu, a president in communist era, won the Romanian Elections in 2000. The task is to get the information that under what circumstance Iliescu won the election, and what his actions toward west and reform in Romania could be.

1. What is your level of knowledge about this story?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

2. Do you recall reading a newspaper article (either in print or online), or listened to or watched a news clip about this story?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

3. What is your level of knowledge about Romania?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

4. What is your level of knowledge about communist era?

	1	2	3	4	5	6	7	8	9	10	
No knowledge at all											Expert

Figure 41: Topic 4 - Background Knowledge Questionnaire

C.4.1 Background Knowledge

C.4.2 Task Description

Iliescu, a president in communist era, won the Romanian Elections in 2000. The task is to get the information that under what circumstance Iliescu won the election, and what his actions toward west and reform in Romania could be.

From the articles, find snippets of relevant text to the following questions:

1. **What is the biography of Iliescu?**
2. **How many voters are there?**
3. **What is the percentage of wining?**
4. **What is the role of Tudor?**
5. **What is the Iliescu's comments about market economy, joining EU and NATO?**

And here is the list of questions you colleagues have worked on:

1. What is the role of Iliescu?

2. What is the role of previous communist rule?
3. How many votes are counted?
4. When is the election?
5. What is the role of the election?

APPENDIX D

SUBJECTIVE QUESTIONNAIRE

D.1 SESSION WITH NO SNS

	Not at all		Somewhat		Extremely
1. Was it easy to find relevant documents with this system?	1	2	3	4	5
2. Was it easy to find relevant *passages* within the retrieved documents?	1	2	3	4	5
3. Did you have a positive experience searching for relevant documents with this system?	1	2	3	4	5
4. Did you feel the system's output was sufficient to help you answer the question?	1	2	3	4	5
5. Did you have enough time to find all the relevant information you were looking for?	1	2	3	4	5
6. Would it be easier to find relevant passages and documents knowing about which documents other users have looked at and collected information from?	1	2	3	4	5

Rank the followings in terms of their importance on helping you to select a document (1 is the highest)

☐ Ranking of the document in the search result

☐ Abstract of the document

Figure 42: Subjective Questionnaire - Sessions with No SNS

D.2 SESSION WITH SNS

	Not at all		Somewhat		Extremely
1. Was it easy to find relevant documents with this system?	1	2	3	4	5
2. Was it easy to find relevant *passages* within the retrieved documents?	1	2	3	4	5
3. Did you have a positive experience searching for relevant documents with this system?	1	2	3	4	5
4. Did you feel the system's output was sufficient to help you answer the question?	1	2	3	4	5
5. Did you have enough time to find all the relevant information you were looking for?	1	2	3	4	5
6. The human icon was used to specify documents visited by other users and the filling level represented the number of previous visits (More filling for higher number of visits). Did you find the human icon meaningful?	1	2	3	4	5
7. The highlight icon was used to specify documents highlighted by other users and the filling level of the highlights represented the number of highlights (More filling for higher number of highlights). Did you find the highlight icon meaningful?	1	2	3	4	5
8. Did you find it useful to know what documents were selected by other users?	1	2	3	4	5
9. Did you find it useful to know what documents were highlighted by other users?	1	2	3	4	5
10. Did you find it useful to view highlighted passages by other users?	1	2	3	4	5
11. Did you find it useful to know the number of times each document was visited by previous users?	1	2	3	4	5
12. Did you find it useful to know the number of times each document was highlighted by previous users?	1	2	3	4	5
13. The tables on top of the page was designed to facilitate navigating to documents highlighted or visited by other users. Did you find it useful?	1	2	3	4	5

Rank the followings in terms of their importance on helping you to select a document (1 is the highest)

- ☐ Ranking of the document in the search result
- ☐ Abstract of the document
- ☐ Being visited by others
- ☐ Being highlighted by others

Figure 43: Subjective Questionnaire - Sessions with SNS

BIBLIOGRAPHY

[1] J. Ahn, P. Brusilovsky, and R. Farzan. Investigating users' needs and behaviors for social search. In *Workshop on New Technologies for Personalized Information Access (PIA 2005). In conjunction with International Conference on User Modeling*, 2005.

[2] J. Ahn, P. Brusilovsky, D. He, J. Grady, and Q. Li. Personalized web exploration with task models. In *Proceedings of the 17th international conference on World Wide Web*, 2008.

[3] E. Bonabeau, M. Dorigo, and G. Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, 1999.

[4] A. Broder. A taxonomy of web search. In *SIGIR Forum*, volume 36, pages 3–10, 2002.

[5] P. Brusilovsky, G. Chavan, and R. Farzan. Social adaptive navigation support for open corpus electronic textbooks. In *Proceedings of Third International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pages 24–33, 2004.

[6] G. Buscher, A. Dengel, and L. van Elst. Query expansion using gaze-based feedback on the subdocument level. In *Proceedings SIGIR conference on Research and development in information retrieval*. ACM Press, 2008.

[7] M. Chalmers, A. Dieberger, K. Höök, and Å. Rudström. Social navigation and seamful design. *Cognitive Studies*, 11(3):1–11, 2004.

[8] C. H. Chen, W. H. Chang, and W. T. Chang. Gender differences in relation to wayfinding strategies, navigational support design, and wayfinding task difficulty. *Journal of Environmental Psychology*, 2008.

[9] E. H. Chi, M. Gumbrecht, and L. Hong. Visual foraging of highlighted text: An eye-tracking study. In *Proceeding of Human-Computer Interaction (HCII'07)*, volume 4552, pages 589–598. Springer-Verlag, 2007.

[10] M. Claypool, P. Le, M. Waseda, and D. Brown. Implicit interest indicators. In *Proceedings of ACM Intelligent User Interfaces*, pages 33–40, anta Fe, New Mexico, USA, 2001.

[11] E. Cutrell and Z. Guan. What are you looking for? an eye-tracking study of information usage in web search. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2007.

[12] B. J. Dahlen, J. A. Konstan, J. Herlocker, N. Good, A. Borchers, and J. Riedl. Jump-starting movielens: User benefits of starting a collaborative filtering system with "dead data". In *University of Minnesota TR 98-017*, 1998.

[13] A. Dieberger. Supporting social navigation on the world wide web. *International Journal of Human-Computer Interaction*, 46:805–825, 1997.

[14] A. Dieberger. Where did all the people go? a collaborative web space with social navigation information. In *Poster at WWW9*, The Hague, The Netherlands, May 2000.

[15] A. Dieberger, K. Höök, M. Svensson, and P. Lönnqvist. Social navigation research agenda. In *Proceeding of CHI extended abstracts on Human factors in computing systems*, 2001.

[16] P. Dourish and M. Chalmers. Running out of space: Models of information navigation. In *Proceedings of HCI'94*, 1994.

[17] J. Dron. Termites in the schoolhouse: Stigmergy and transactional distance in an e-learning environment. In L. Cantoni and C. McLoughlin, editors, *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, pages 263–269, Chesapeake, VA, 2004. AACE.

[18] T. Erickson and W. A. Kellogg. *Designing Information Spaces: The Social Navigation Approach*, chapter Social Translucence: Using Minimalist Visualizations of Social Activity to Support Collective Interaction, pages 17–42. Springer, 2003.

[19] J. P. Fan and R. D. Macrediel. *Advances in Web-based Education: Personalized Learning Environments*, chapter Gender Differences and Hypermedia Navigation: Principles for Adaptive Hypermedia Learning Systems. Idea Group Inc, 2006.

[20] R. Farzan and P. Brusilovsky. Social navigation support in e-learning: What are the real footprints? In *Proceedings of Third Workshop on Intelligent Techniques for Web Personalization. At Nineteenth International Joint Conference on Artificial Intelligence*, 2005.

[21] R. Farzan and P. Brusilovsky. Social navigation support through annotation-based group modeling. In *Proceedings of 10th International User Modeling Conference*, pages 463–472. Berlin: Springer Verlag, 2005.

[22] R. Farzan and P. Brusilovsky. Annotated: A social navigation and annotation service for web-based educational resources. *New Review in Hypermedia and Multimedia*, 14(1):3–32, 2008.

[23] R. Farzan and P. Brusilovsky. Where did the researchers go? supporting social navigation at a large academic conference. In *Proceedings of Hypertext 2008*, 2008.

[24] R. Farzan, M. Coyle, J. Freyne, P. Brusilovsky, and B. Smyth. Assist: adaptive social support for information space traversal. In *Proceedings of 18th conference on Hypertext and hypermedia*, pages 199–208. ACM Press, 2007.

[25] N. Ford and D. Miller. Gender differences in internet perceptions and use. In *Proceedings of Aslib*, volume 48, pages 183–192, 1996.

[26] N. Ford, D. Miller, and N. Moss. The role of individual differences in internet searching: An empirical study. *Journal of the American Society for Information Science and Technology*, 52(12):1049–1066, 2001.

[27] J. Freyne, R. Farzan, P. Brusilovsky, B. Smyth, and M. Coyle. Collecting community wisdom: Integrating social search and social navigation. In *Proceedings of International Conference on Intelligent User Interfaces*, pages 52–61, Honolulu, HI, USA. ACM Press.

[28] J. Golbeck. Trust on the world wide web: A survey. *Foundations and Trends in Web Science*, 1(2), 2008.

[29] J. H. Goldberg and X. P. Kotval. Computer interface evaluation using eye-movement: methods and constructs - its psychological foundation and relevance to display design. *International Journal of Industrial Ergomics*, 1999.

[30] D. He, P. Brusilovsky, J. W. Ahn, J. Grady, R. Farzan, Y. Peng, Y. Yang, and M. Rogati. An evaluation of adaptive filtering in the context of realistic task-based information exploration. *Information Processing and Management*, 44(2):511–533, 2008.

[31] W. Hill, L. Stead, M. Rosenstein, and G. W. Furnas. Recommending and evaluating choices in a virtual community of use. In *CHI Conference on human factors in computing systems*, pages 194–201, 1995.

[32] W. C. Hill, J. D. Hollan, D. Wroblewski, and T. Mccandless. Edit wear and read wear. In P. Bauersfeld, J. Bennett, , and G. Lynch, editors, *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 3–9, 1992.

[33] K. Höök, D. Benyon, and A. J. Munro. *Designing Information Spaces: The Social Navigation Approach*. Berlin: Springer, 2003.

[34] K. Höök, N. Dahlbck, and M. Sjölinder. Individual differences and navigation in hypermedia. In *Proceedings of ECCE-8*, pages 195–200, 1996.

[35] R. J. K. Jacob and K. S. Karn. Eye tracking in human computer interaction and usability research: ready to deliver the promises (section commentary). *The minds eye: cognitive and applied aspects of eye movement research*, pages 573–605, 2003.

[36] T. Joachims, L. Granka, and B. Pan. Accurately interpreting clickthrough data as implicit feedback. In *SIGIR'05*, 2005.

[37] P. B. Kantor, E. Boros, B. Melamed, V. Mekov, B. Shapira, and D. Neu. Capturing human intelligence in the net. In A. Press, editor, *Communications of the ACM*, volume 43, pages 112–-116, 2000.

[38] J. Kurhila, M. Miettinen, P. Nokelainen, and H. Tirri. Educo - a collaborative learning environment based on social navigation. In Springer, editor, *Proceedings of 2nd International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pages 242–252, 2002.

[39] C. A. Lawton. Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles*, 30(11/12), 1994.

[40] Y. Liu and K. M. Passino. Swarm intelligence: Literature overview. The Ohio State University, 2000.

[41] L. Lorigo, B. Pan, H. Hembrooke, T. Joachims, L. Granka, and G. Gay. The influence of task and gender on search and evaluation behavior using google. *Information Processing and Management*, 42:1123–1131, 2006.

[42] D. R. Millen and J. Feinberg. Using social tagging to improve social navigation. In *Workshop on the Social Navigation and Community based Adaptation Technologies*, 2006.

[43] D. R. Millen, J. Feinberg, and B. Kerr. Dogear: Social bookmarkin in the enterprise. In *Proceedings of CHI*, 2006.

[44] T. Mooradian, B. Renzl, and K. Matzler. Who trusts? personality, trust and knowledge sharing. *Management Learning*, 37(4):523–540, 2006.

[45] M. O'Brien and M. T. Keane. Modeling result-list searching in the world wide web: The role of relevance topologies and trust bias. In *Proceedings of the Twenty-Eight Annual Conference of the Cognitive Science Society*, 2006.

[46] W. Pan. Akaike's information criterion in generalized estimating equations. *Journal of Biometrics*, 57:120–125, 2001.

[47] P. Pirolli. An elementary social information foraging model. In *Proceeding of CHI'09*. ACM Press, 2009.

[48] P. Pirolli and S. K. Card. Information foraging. *Psychlogical Review*, 1999.

[49] A. Poole and L. J. Ball. *Encyclopedia of Human Computer Interaction*, chapter Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. Information Science Reference, 2006.

[50] K. Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124:372–422, 1998.

[51] W. R. Reader and S. J. Payne. Allocating time across multiple texts: Sampling and satisficing. *Human Computer Interaction*, 22:263–298, 2007.

[52] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: An open architecture for collaborative filtering of netnews. In *CSCW*, 1994.

[53] Å. Rudström and P. Fagerberg. Socially enhanced travel booking: a case study. *Journal of Information Technology and Tourism, special issue on Travel Recommender Systems*, 6(3), 2004.

[54] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen. *The Adaptive Web: Methods and Strategies of Web Personalization*, volume 4321 of *Lecture Notes in Computer Science*, chapter Collaborative Filtering Recommender Systems. Springer, 2007.

[55] A. I. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock. Methods and metrics for cold-start recommendations. In *SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2002.

[56] B. Smyth, E. Balfe, J. Freyne, P. Briggs, M. Coyle, and O. Boydell. Exploiting query repetition and regularity in an adaptive community-based web search engine. *User Modeling and User-Adaptated Interaction*, 14(5):383–423, 2004.

[57] M. Svensson, K. Höök, and R. Coster. Designing and evaluating kalas: a social navigation system for food recipes. *ACM Transactions on Computer-Human Interaction*, 12(3):374–400, 2005.

[58] M. Svensson, K. Höök, J. Laaksolahti, and A. Waern. Social navigation of food recipes. In *Proceedings of CHI*, Seattle, April 2001. ACM.

[59] M. Svensson, J. Laaksolahti, K. Höök, and A. Laaksolahti. A recipe based on-line food store. In *Proceedings of 5th international Conference on intelligent User interfaces*, pages 260–263. ACM Press, 2000.

[60] M. W. H. Weenig and M. Maarleveld. The impact of time constraint on information search strategies in complex choice tasks. *Journal of Economic Psychology*, 23(6):689–702, December 2002.

[61] A. Wexelblat and P. Maes. Footprints: History-rich tools for information foraging. In *Proceeding of ACM Conference on Human-Computer Interaction (CHI'99)*, pages 270–277. ACM Press, 1999.

[62] S. Wilkinson and S. Payne. Eye tracking to identify strategies used by readers seeking information from on-line texts. In *Proceedings of Eurpoean conference on Cognitive*

ergonomics: trust and control in complex socio-technical systems, pages 115–116. ACM, 2006.

[63] J. Wu and K. Aberer. Swarm intelligent surfing in the web. In *Proceeding of Web Engineering*, pages 277–284. Springer Berlin, 2003.

[64] H. B. Zur and S. J. Breznitz. The effects of time pressure on risky choice behavior. *Acta Psychologica*, 47:89–104, 1981.