Social Annotations in Web Search

Aditi Muralidharan
Computer Science Department
UC Berkeley
Berkeley, CA
aditi@cs.berkeley.edu

Zoltan Gyongyi
Google, Inc.
Mountain View, CA
zoltang@google.com

Ed H. Chi
Google, Inc.
Mountain View, CA
edchi@google.com

ABSTRACT

We ask how to best present social annotations on search results, and attempt to find an answer through mixed-method eye-tracking and interview experiments. Current practice is anchored on the assumption that faces and names draw attention; the same presentation format is used independently of the social connection strength and the search query topic. The key findings of our experiments indicate room for improvement. First, only certain social contacts are useful sources of information, depending on the search topic. Second, faces lose their well-documented power to draw attention when rendered small as part of a social search result annotation. Third, and perhaps most surprisingly, social annotations go largely unnoticed by users in general due to selective, structured visual parsing behaviors specific to search result pages. We conclude by recommending improvements to the design and content of social annotations to make them more noticeable and useful.

Author Keywords

social search; information seeking; eye tracking

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces—*Graphical User Interfaces*

INTRODUCTION

When searching for a restaurant in New York, does knowing that aunt Carol tweeted about *Le Bernardin* make it a more useful result?

In real life, all of us rely on information from our social circles to make educated decisions, to discover new things, to stay abreast of news and gossip. Similarly, the quantity, diversity, and personal relevance of social information online makes it a prime source of signals that could improve the experience of our ubiquitous search tasks [8]. Search engines have started incorporating social information, most prevalently by making relevant social activity explicitly visible to searchers through *social annotations*. For instance, Figure 1 shows social annotations as they appeared on Google and Bing search results in early September 2011.
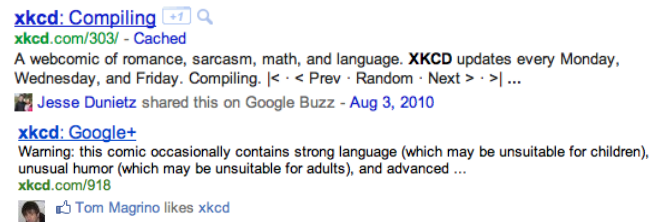
Figure 1. Social Annotations in web search as of sept. 2011 in Google (top), and Bing (bottom).

Annotations are added to normal web search results to provide social context. The presence of an annotation indicates that the particular result web page was shared or created by one of the searcher's online contacts. For example, suppose that a friend of Aunt Carol on Google+ or Facebook searched for "Le Bernardin", and Aunt Carol had previously posted a restaurant review on *Le Bernardin's* Yelp page. If one of the search results happened to be that same yelp.com restaurant page, the searcher would see an annotation like in Figure 1 on that result, explaining that Aunt Carol had reviewed it on some date.

To enable social results in Google search, users have to be signed in and connect their accounts on various social sites to their Google profiles. For social search on Bing, users have to be signed in with their Facebook accounts. After enabling social search, users will occasionally see social annotations on some results when they search the web.

Social annotations so far have generally a consistent presentation: they combine the profile picture and the name of the sharing contact with information about when and where the sharing happened. The annotation is rendered in a single line below the snippet. While there is some apparent convergence in current practice, we know of no research on what content and presentation make social annotations most useful in search for users.

In this paper, we set out to answer this particular question, equipped with intuitions derived from past research on search behavior and perceptual psychology. Intuitively, one could argue that social annotations should be noticeable and attention-grabbing: it is well-documented that faces and face-like symbols capture attention [41, 25, 19], even when users are not looking for them and do not even expect them [27]. Moreover, social annotations should also be useful: users collaborate on search tasks with each other by looking over each others' shoulders and by suggesting keywords [30, 12], so marking results that others have found useful should help.

However, it is not immediately clear how valid these intuitions are in the domain of web search. This particular domain

is known for specialized user behaviors, and it is not obvious how adding social cues might affect them. In particular, users pay almost exclusive attention to just the first few results on a page [23, 16]. It seems that they have learned to trust search engines, because this behavior persists even when the top few results are manipulated to be less relevant [34, 17]. So how does adding an extra layer of social cues in the form of social annotations affect the search process? Are their friends' names and faces powerful enough to draw users attention? Or do learned result-scanning behaviors still dominate?

We designed and conducted two consecutive experiments, which we describe in detail in two major sections below. The first experiment investigated how annotations interacted with the search process in general, while the second focused on how various visual presentations affected the visibility of annotations. In our first experiment, we found that our intuitions were not quite accurate: annotations are neither universally noticeable nor consistently useful to users. This experiment unveiled a colorful and nuanced picture of how users interact with socially-annotated search results while performing search tasks. In our second experiment we uncovered evidence that strong, general reading habits particular to search result pages seem to govern annotation visibility.

In the next section, we give an overview of related research in this area.

RELATED WORK
The literature that is perhaps most immediately connected to our research questions addresses the use of the "social information trail" in web search. Yet, to the best of our knowledge, the work in this area has focused on behind-the-scenes applications, not on presentation or content. Many have proposed ways to use likes, shares, reviews, +1's and social bookmarks to personalize search result rankings [4, 43, 20, 44, 6], but there is comparatively little on making this information explicitly visible to users during web search.

More generally, studies of the information seeking process can help us understand what types of social information could be useful to searchers, and when. Several proposed models capture how people use social information and collaborate with each other during search. Evans and Chi [12] found that people exchanged information with others before, during and after search. Pirolli [36] modeled the cost-benefit effects of social group diversity and structure in cooperative information foraging [37]. Golovchinsky and colleagues [15, 14] categorized social information seeking behaviors in according to the degree of shared intent, search depth, and the concurrency and location of searches.

Social Feedback during Search
As mentioned before, searchers look to other people for help before, during and after web search [12]. They have to resort to over-the-shoulder searching, e-mails, and other ad-hoc tools because of the lack of built-in support for collaboration [30]. Researchers have responded with a variety systems that aid collaboration. These systems share histories of search keyword choices, communicate by voice or text, and share entire search sessions, starting from search keyword combinations all the way to the final web pages that returned the best information [32, 31, 3, 35]. Social bookmarks have also been investigated as a source of helpful signals during search [24].

Social Question-Answering
In social question-answering systems, topical expertise is important to the overall system design. People's knowledge on various topics is indexed, and used to route appropriate questions to them [1, 2, 22]. Evans, Kairam and Pirolli [13] studied social question answering for an exploratory learning task, and found expertise to be an important factor in peoples' decisions to ask each other for information. In real-world question-answering situations, Borgatti and Cross [5] studied organizational learning and identified the perceived expertise of a person as one of the factors that influenced whether to ask them a question. Further bearing out the importance of expertise, Nelson et. al. [33] found that expert annotations improved learning scores for exploratory learning tasks.

There are also suggestions that certain topics such as restaurants, travel, local services, and shopping are more likely to benefit from the input of others, even strangers. For instance, Amazon, Yelp and TripAdvisor have sophisticated review systems for shopping, restaurants and travel respectively. Moreover, Aardvark.com reports that the majority of questions fall into exactly these topic types [22].

Collaborative Activity
Beyond web search and social question-answering systems, simply displaying or visualizing past history can be useful to collaborators. Past research on collaborative work suggests that it is useful to show users a history of activity around shared documents and shared spaces. One of the first ideas in this area was the concept of read wear and edit wear [21] — visually displaying a documents reading and editing history as part of the document itself. A document's read and edit wear is a kind of social signal. It communicates the sections of the document that other people have found worth reading or worth editing. Erickson and Kellogg [11, 10] introduced the term "social translucence" to describe the idea of making collective activity visible to users of collaborative systems. They argued that social translucence makes collaboration and communication easier by allowing users to learn by imitation, and take advantage of collective knowledge. They also argued that, when users' actions are made visible to others, social pressures encourage good behavior. Social annotations can be seen as a kind of social translucency.

Methods
In our experiments, we use eye tracking to study user behavior while varying search-result-page design. In the past, Cutrell and Guan analyzed the effects of placing a "target" high-quality link at different search result ranks [9] on various task completion metrics. The effects of "short", "medium", and "long" snippet lengths on search task behavior was also studied by Guan and Cutrell [17].

STUDY 1: PERCEPTIONS OF SOCIAL ANNOTATIONS

Study Goals

The research cited above suggests that social annotations are broadly useful, perhaps especially on certain topics and from certain people. Our objective was to find out how social annotations influence web search behavior. Specifically, we wanted to understand the importance of contact closeness, contact expertise, and search topic in determining whether or not an annotation is useful.

Study Design

We performed a retrospective-interview-based study with N=11 participants. The participants did not know the intent of the study, and were told that we were evaluating a search engine for different types of search tasks.

Eye-tracking data was recorded so that the researcher could observe the participant more closely. The participants' gaze movements were displayed to the researcher on a separate screen, and recorded for replay during the interview. This way, the researcher could monitor whether participants were looking at or ignoring social annotations, and concentrate on those events during the retrospective interview.

A major challenge in this study was to make the study experience as organic as possible. We wanted users to truly reveal their natural search tendencies without suspecting that we were studying social annotations.

Study Procedure

In the first half of the study, participants performed a series of 18-20 search tasks, randomly ordered. Half the search tasks were designed so that search queries would bring up one or two social annotations in the top four or five results. In the second half of the study, immediately after the tasks, participants were interviewed in retrospective think-aloud (RTA) format. They were asked to take the researcher through what they were doing on some of the tasks while watching a replay of a screen capture (including eye traces) of the tasks with the researcher. The researcher would ask probing questions for clarification, and then move on to talk about another task.

If the participant never mentioned the social annotations, even after an interview on all the tasks in which they were present, the interview procedure was slightly different. The researcher would return to a screen capture of one of the tasks in which there was a social annotation. The researcher would point out the annotation explicitly and ask a series of questions, like, "What is that? Did you notice it? Who is this person? Tell me what you think about this." The goals were to find out what the participant thought the annotation was, whether they noticed it at all, and to understand whether or not they perceived it as useful, and why. After the annotation had been discussed, the researcher would revisit the remaining social annotations and repeat the procedure for each of them.

Personalization

From past research [5, 22, 13], we suspected that the expertise of the contact might influence how an annotation is perceived. For example, a participant might react differently to a hiker friend on a result about hiking trails than on a topic on which that friend is completely ignorant. This meant that we would have to design tasks for each participant individually, and could not simply insert fake annotations into a standard set of tasks. Even if the names and faces were personalized individually to be real contacts, there would be no guarantee that the tasks we created would match those contacts' expertise areas. To ensure that we were eliciting representative reactions to social annotations, we designed the search tasks individually for each participant by looking at the links and blog posts shared by the contacts in social networks they had linked to their accounts.

Participants

Every social annotation seen by each participant was real, just like they would have seen outside the lab, with no modifications. This meant that we had to infer, for each participant in advance of the study, which queries would bring up social annotations, and design individual search tasks around these URLs for each participant. In this way, we created 8-10 social search tasks for each participant across different topics (depending on the availability of searchable, annotated URLs).

It was somewhat difficult to find enough participants with these constraints. Our first step was to find participants who were willing to allow enough access to their personal data to personalize the study. We recruited around 60 respondents who gave us this permission. Of these, we found 11 respondents with enough data in their social annotations to see annotations in 10 different searches.

Search Tasks

Half the tasks were designed to bring up social annotations, and the other half were "non-social". The tasks were framed as questions, and not as pre-assigned keyword queries. We did not provide pre-assigned queries because we did not want to raise suspicions that there was something special about the search pages or the particular keyword combination. Some sample search tasks are shown in Table 1. In the case of social tasks, prompts were worded so that many reasonable keyword combinations would bring up one or two social annotations in the top four or five search results.

Topic	Search Task
How-to	How do you make a box kite?
Recipe	How do you make milk bar pie?
Product	Find a good laptop case for your macbook pro.
Local	Find a good sweet shop in Napa, CA.
Entertainment	Is the album "The Suburbs" by Arcade Fire any good?
News	What is going on with stem cells latelyI?
Fact-finding	Where did Parkour originate?
Navigation	What is the website of Time Warner's youtube channel?

Table 1. Samples of search tasks in different topics. Participants were not shown the topic category, they were just given a strip of paper with the task question printed on it. They started each task at the default search engine start page.

We distributed the tasks for each participant across suspected useful and non-useful topics. We thought annotations would
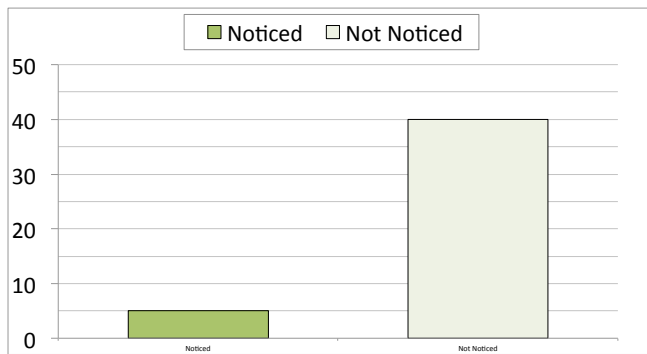
Figure 2. The total number of social annotations that were noticed.

be useful on the topics of product search, local services, how-tos, recipes, news and entertainment. Participants each performed up to 4 tasks in each of these topic categories, of which two had annotations. We thought annotations may not be useful for navigation and fact-finding, so participants performed 2 tasks in each of those categories, one with annotations and one without.

The inclusion of a category was subject to the availability of social annotations in that category. For example, if we could not find any *product search* annotations for a participant, we did not ask them to perform any product search tasks at all. This resulted in minor variation in the number of search tasks. Some participants had more annotations in more categories than others.

Results

To understand the space of participants' responses, we first created an affinity diagram by grouping responses that seemed to be similar in meaning. Once categories of responses emerged, we went through the responses yet again, coding the responses into different categories, and counting the numbers in each category.

We did not analyze the eye tracking data we gathered due to the heavily-personalized and unrestricted nature of the search tasks: each participant saw a different set of pages, and participants were allowed to scroll and to revisit search pages with the back button. This made standardized region-of-interest annotations very difficult to make. We used the eyetracking videos primarily as a memory aid for the participant during the RTA's, but also as a supplement to the interview results, to arrive at the conclusions below.

Social Annotations Go Unnoticed

Figure 2 shows that most social annotations were not noticed by participants. Of the 45 annotations in our experiment that appeared above the page fold, 40 (89%) were not noticed. By "not noticed", we mean that participants explicitly said, when the annotation was pointed out, that they did not notice the annotation while they were searching. Often, the participants were surprised they had missed it because it was from a close friend, co-worker, old friend, or boss. In fact, of the 40 missed annotations, 35 were from people the participants

recognized as friends, co-workers, family members, or acquaintances (but 5 were from people the participants did not recognize).

Reactions to Seen Annotations

The following summary of our participants reactions reveals concerns about privacy and objectivity, as well the value of closer ties and known expertise.

Participant 5 (P5) noticed annotations on two successive queries and commented on them unprompted while searching. She highlighted the first annotation she saw with her cursor and exclaimed that the annotation was "creepy" even though the annotation was from "a friend."

> "... this friend is a friend, but the algorithm or whatever makes her show up there doesn't know how close we are, or how much I respect her opinion... "

The next query, however, was a much more positive experience. This time, the annotation was from her personal trainer on a fitness-related topic. This annotation was better because she trusted her trainer's expertise on the subject, and would actually turn to him for advice in real life.

The other participants did not comment on social annotations while searching, but had informative reactions during the RTA. P11 explained why he clicked on a Yelp result on which he noticed a friends name and face:

> "Yeah, I directly know her, it's not just somebody, like a friend of a friend... "

Finally, P4 and P6 had both seen social annotations outside the lab, but did not click on the annotations even though they saw them. They gave different reasons why. P6 passed over a friend-annotated blog post, and instead chose Wikipedia because:

> "The immediate thing I thought was that he [the friend] edits Wikipedia pages, he's been doing it for a long time"

P4, however, gave a different reason for not clicking on an annotated search result:

> "I don't necessarily want to see what they've looked at, I want to see sources that I think are credible..."

Annotations Are Useful for "Social" and "Subjective" Topics

Ten out of the eleven participants mentioned that social annotations would be useful on topics like restaurants, shopping, searching for businesses, shopping for expensive items, and planning events for other people. The remaining participant did not specify a topic. When asked to generalize, participants used the words "social", "subjective" or "reviews" to describe the category of topics. When asked why the information would be useful, participants said that it would be good extra information in decision-making if they knew the person had good taste, was knowledgeable, or was trusted by them to have good information on the topic.

Another category of useful topics, brought up by 7 out of the 11 participants, is best described as personal, or hobby-related. It seemed that participants would be curious to see

close contacts or family members in social annotations, so that they could connect with them over a shared hobby or activity, or because it might be nice to discover that they had a shared interest.

Annotations from Knowledgeable Contacts Are Useful
All participants said that annotations would be useful when they came from people whom searchers believed had good taste, knowledge or experience with the topic, or had similar likes and dislikes. This was not restricted to contacts they knew personally, as celebrities, bloggers, or respected authorities on a topic were also indicated to be useful.

Closer Relationships Make Annotations More Useful
Nine out the eleven participants said that annotations from strong-tie contacts (such as close friends, in regular contact, or family members) would be more useful than more distant contacts. Four participants made the distinction between interestingness and relevance. To paraphrase their responses, annotation from a very close friend might be interesting because of the information it gave about the friend's interests or activities, but it may not provide any relevant useful information about the quality of the result, or make it any easier to complete the task.

Seven out of the eight participants who saw annotations from strangers indicated that they would ignore those annotations. This included people they did not recognize, and people they did not have a significant relationship with. One participant said that he would be confused, and would want to know what his relationships to that person was. When asked whether seeing strangers was a negative, or simply irrelevant, 7 participants responded that it was irrelevant.

The Searcher's Mindset Could Affect Usefulness
Three out of the eleven participants explicitly mentioned that they would only click on the social annotation or talk to their friend later on about seeing the social annotations if they had time, or were simply exploring a topic space. The remaining participants did not specify when they would click or follow-up with a friend on a social annotation.

Discussion
This study revealed a counter-intuitive result. Despite having the names and faces of familiar people, and despite being intended to be noticeable to searchers, subjects for the most part did not pay attention to the social annotations.

Our questions about contact closeness, expertise, and topic were answered by the reactions captured during the *retrospective* interviews. These interviews revealed the importance of contact expertise and closeness, and the importance of the search topics in determining whether social signals are useful, thus echoing pas findings on the role of expertise in social search [13].The interviews also provided some high-level understanding of the ways that people use, and want to use, social information during web search.

Nevertheless, a challenge emerged, which is our need to understand the lack of attention to social annotations, and finding ways to improve their presentation. Having confirmed that many types of social information could indeed be useful to searchers, we had to ask why the annotations conveying this information were largely ignored. In the next section, we describe the follow-up experiment we conducted to get to the bottom of this mystery.

STUDY 2: PRESENTATION OF SOCIAL ANNOTATIONS
Past work shows that people discriminate between the different parts of a search result and do not linearly scan pages from top to bottom. Titles, URLs, and snippets receive different amounts of attention [9], and, in the sample of over 600 gaze paths analyzed in [26], 50% contained regressions to higher-ranked results and skips to lower-ranked results.

Study Goals
As stated above, the goal of our second study was to find out why so many of the social annotations in the first study went unnoticed. We designed an experiment to investigate what would happen to users' page-reading patterns when social annotations were added in various different design variations. We hoped to find behaviors anchored in the familiar presentation of search results that would explain why the social annotations in the first experiment were ignored. Our particular research questions were:

1. Will increasing the sizes of the profile pictures make social annotations more noticeable?

2. Are there learned reading behaviors that prevent participants from paying attention to social annotations?

Study Design
We performed a mixed-method eye-tracking and retrospective-interview-based study with N=12 participants. The participants did not know the intent of the study, and were told that we were evaluating a search engine for different types of search tasks.

Participants
We recruited 15 non-computer-programmers from within our organization, but had to discard data from 3 of them due to interference with their eye glasses. As our second study focused on presentation issues only, we decided on less-personalized annotations than in the first study. Accordingly, we did not have to analyze participants private data, and thus ended up with a simpler recruiting process.

Personalization
In order to control the stimuli presented to participants, we did not personalize the search tasks. We used the same set of tasks across all participants.

The only personalization was the names and faces of people in the annotations. These were participants' real co-workers, but annotations appeared on results of our choosing, and not results that had really been shared by those people (in contrast with the first study). Using our internal employee directory
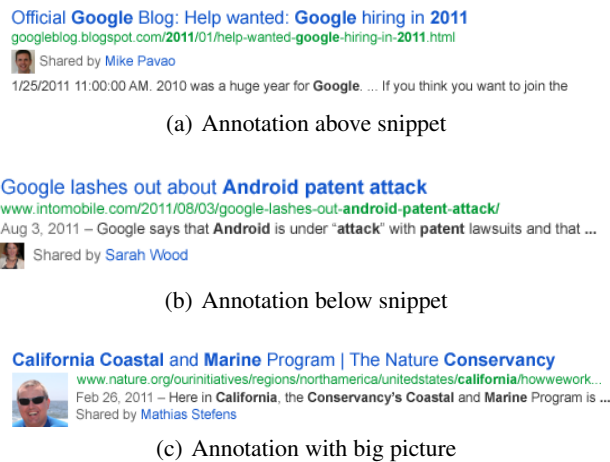
(a) Annotation above snippet

(b) Annotation below snippet

(c) Annotation with big picture

Figure 3. The different annotation variations in Study 2.

and organizational chart, we found the names and pictures of 18 people who were either on the same team as the participants or had their desks within 10 meters of them, reasoning that the participant would recognize most of these people. Their names and faces were then pasted into the static mock-ups of web pages with social annotations.

Social Annotation Design Variations

Our goal was to see whether changing snippet length, annotation placement, and picture size changed the amount of attention (measured in number of fixations) given to the annotation. For snippets, we had 1-line, 2-line, and 4-line snippets. The annotation's presentation within the result was varied to be either above the snippet or below the snippet (Figure 3a–b). The annotated result was either the first result on the page or the second result. Additionally, to test our hypothesis about the faces in the annotations being too small to be noticed, we added another annotation presentation condition by using a $50{\times}50$ picture placed in-line with the snippet, as shown in Figure 3c.

Together these annotation variations, snippet length variations, and result position variations created a 3x3x2 = 18 different conditions, as follows: [big picture inline, small picture below snippet, small picture above snippet] $\times$ [1, 2, 4] line snippets $\times$ [1st or 2nd search result]. These variations were interleaved with an equal number of baseline non-annotated result pages, bringing the total to 36 tasks.

Stimuli

Participants viewed and clicked on 36 mock-ups of search result pages. Half of these had social annotations, and half did not, and the social and non-social pages were interleaved.

The motivation for using both annotated and non-annotated mock-ups was twofold. First, we wanted to avoid raising suspicions about the nature of the study, and second, the non-annotated pages provided identical baselines on which we could compare all participants.

The social pages were generated with an image editor. We generated pages with different snippet lengths, annotation positions, and picture sizes. Then, we pasted in the names and faces of office-mates and team-mates to personalize the mock-ups for each participant. The non-social pages were generated by taking screenshots of search results.

Procedure

Participant Conditions
Due to their prominent size, we suspected that the big $50{\times}50$px pictures might prime the participants to the social nature of the experiment. We therefore divided the participants into two conditions to avoid an undetected priming bias: the first group (N=5, 2 more discarded) saw the big-picture variants first, before any other type of annotation, and the second group (N=7, 1 more discarded) saw the big-picture variants last, only after they had seen all the $21{\times}21$px variants.

Study Procedure
In the first part of the study, participants performed 36 consecutive search tasks. For each task, they were first shown a screen with a task prompt, and asked to imagine having that task in mind. Once they had read the task prompt, they pressed the space bar. This took them to the search page mock-up. They were instructed to view the page as they would normally, and to a click on a result in the end.

In the second part, the participants were retrospectively interviewed about some of the search tasks. The researcher played back a screen capture of their eye movements, and asked questions. Unlike the first study, the interviews were short. We directly asked whether they had noticed the annotation, who the person was, and which annotation presentation they preferred: above-snippet, below-snippet, or big picture.

Results

For all participants, we measured how the number of fixations on annotations varied with snippet length, annotation placement, and annotated result. In addition to the results reported in the following sections, we performed a linear regression, controlling for between-participant variation, and picture order. Further, for succinct visual evidence, we have supplemented some of the quantitative results below with gaze maps averaged across all the participants.

Our results are the same when analyzed using fixation count or fixation duration. We chose fixation count as the presented metric because it is more intuitive to think about whether users actually moved their eyes to the annotations.

Longer Snippets Lead to Less Time on Annotations
The graph in figure 4 shows the average number of fixations on various elements of the search result item, compared across different snippet lengths. We can see that annotations below a 1-line snippet get almost twice as many fixations compared to annotations below a 4-line snippet. In our linear regression, 2-line and 4-line snippets received negative coefficients (after controlling for between-participant variation and picture-size order), meaning that they decreased the fixation count, with
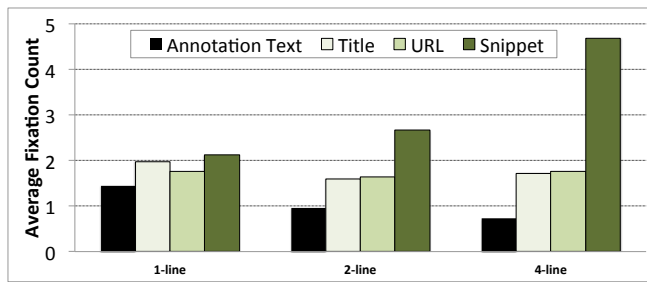
Figure 4. Average fixation count on various result elements vs. the length of the snippet (N=12). We showed participants 3 snippet lengths: 1-line, 2-line and 4-line. Fixation counts for the annotation are drawn in black, values for snippet, title, and URL are in shades of green.

Figure 5.   The snippet-length effect shown for the annotation-below-snippet condition. The different lengths were 1-line (top) 2-line (middle) and 4-line (bottom). These averaged gaze maps show that the longer the snippet, the fewer the fixations on the annotation.

$(b = -0.51, t(203) = -1.85, p < 0.07)$ and $(b = -0.74, t(203) = -2.56, p < 0.01)$ respectively. An example of this effect for the below-snippet presentation is shown in the averaged gaze heat-maps in figure 5. It is clearly visible that, on average, the longer the snippet above the annotation, the fewer fixations it got (darker regions correspond to fewer fixations).

The effects of snippet length on the other result elements are in line with past findings [9]. Fixations to the snippet increase with snippet length, and fixations to URL and title are relatively constant.

Bigger Pictures Get More Attention
As one might expect intuitively, the 50×50 pictures have a dramatically larger average number of fixations than the smaller 21×21 pictures. Figure 6 shows the effect on number of fixations to pictures. The critical threshold here is a value of 1. A value above 1 means that the element, on average, receives attention. A value below 1 means the opposite. Figure 6 shows that the big pictures receive around 1.3 fixations on average, but the small pictures only receive 0.1. Not surprisingly, the conclusion therefore is that big pictures of faces get noticed, whereas small ones generally do not.
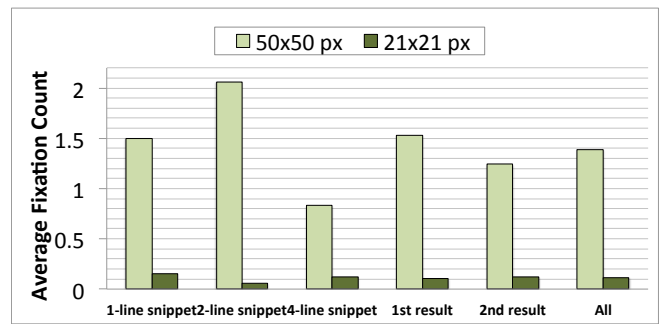

Figure 6. Average fixation count on the pictures in the annotations vs. the size of the picture, for each of the different presentation variations (N=12). We showed participants two sizes, 50x50px and 21x21px.
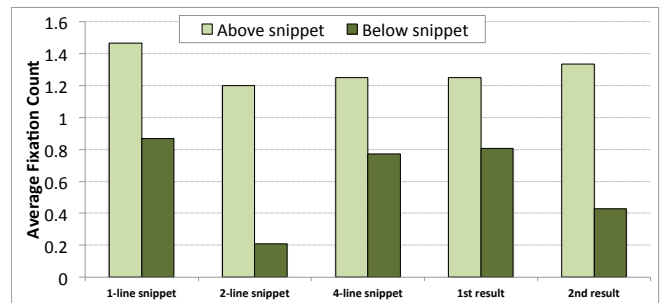

Figure 7. Average fixation count on annotations vs. annotation placement, for each of the different presentation variations (N=12). Annotations were placed either above the snippet or below the snippet.

The effect was significant in our linear regression. The large-picture condition received a positive coefficient $(b = 0.72, t(203) = 2.62, p < 0.01)$, meaning that increasing the picture size increased the number of fixations .

Annotations Above the Snippet Get More Attention
Annotations above the snippet get uniformly more fixations than annotations below the snippet. The graph in figure 7 shows that the effect is true for all snippet lengths and result positions. The above-snippet condition received a positive coefficient in our linear regression $(b = 0.59, t(203) = 2.03, p < 0.04)$, meaning that annotations above the snippet got more fixations.

The heat map in Figure 8 shows an example of the placement effect in the 4-line-snippet condition. It is obvious from the figure that the annotation got more fixations (brighter region) when it was placed above the snippet.

Result 1 Gets More Attention Than Result 2
Figure 9 shows the effect of result position on attention to annotations, averaged across all annotation types, snippet lengths, and picture sizes. Annotations on the first result receive about 1.3 fixations on average, but annotations on the second only receive 0.8 fixations on average.

In our regression model, the 2nd-position condition received a negative coefficient $(b = -0.62, t(203) = -2.66, p < 0.01)$, meaning that it reduced the number of fixations.

Figure 8. The annotation-placement effect, shown for the 4-line snippet condition. These averaged gaze maps show that annotations that were placed above the snippet (top) got more fixations than those placed below (bottom).
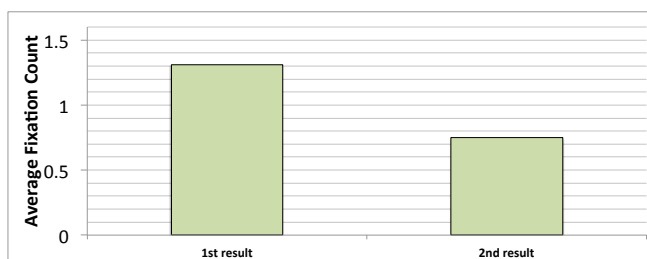
Figure 9. Average fixation count on annotation vs. result position, averaged over all the presentation variations. Annotations were either on the first result or the second result.

Statistical Disclaimer

The experiment produced only one data point per participant for each configuration of annotated result, placement, and snippet length, giving N = 12×18 fixation-count measurements. Therefore, we fit a simple additive model to find the effects of each variable on fixation. The additive model is a crude approximation, and the data are non-normal, so our *p*-values should only be interpreted as a rough guide to statistical significance.

DISCUSSION

In the first study, participants were often surprised when social annotations were pointed out to them. From their comments, they seemed to believe they did not notice the annotations because they were engrossed in their search tasks.

In our second study, we found that (1) users always paid attention to the URLs and titles, and increased their attention to the social annotations when (2) the summary snippets were shorter, (3) pictures were bigger, and (4) when the annotations were placed above the snippet summary.

Together with past research on search-page reading habits, the second study's results suggest that users perform a *structural parse*: they break the page down into meaningful structures like titles, urls, snippets, etc. and only pay attention to certain elements within the structure. This in turn implies that users' blindness to annotations might be caused by learned

lack of attention to anything other than titles, URLs and snippets. Users are so focused on performing their task, and social annotations are not part of their existing page-reading habits, so they simply skip over them and act as if they are not there.

The phenomenon of lack of attention causing functional blindness to clearly visible stimuli has been documented with many different types of activities. Pilots have failed to see another plane blocking a landing runway [18] and spectators of dodgeball have failed to notice a person in a gorilla suit walking across the playing field [40]. Mack, Rock, and colleagues [38, 28, 27], studied the phenomenon extensively, and gave it the name "inattentional blindness".

The bigger profile pictures, however, drew attention as expected from studies of attention capture by faces [41, 25, 19, 27]. So, if the pictures in the first study had been bigger, the annotations might have been noticed more. At a small size however, they were not capable of disrupting users' page scanning patterns.

In search result pages, titles stand out with their blue color and large font, urls stand out in green, and matching keywords are marked out in bold text. Human beings have a cognitive bias that leads us to learn and remember information that is visually prominent [42]. Highlighted or underlined text is remembered and learned better than normal text, even if the highlights are not useful [39]. Highlighted text is also given increased visual attention as measured by an eyetracker [7]. The observed selective attention to certain elements might stem from this effect, combined with learning over time.

However, the results also suggest that we can direct more attention towards a social annotations by manipulating page structure to our advantage. Attention can be gained by placing annotations above the snippet, shortening the snippet, and increasing annotation picture size.

While we can manipulate the visual design to make annotations more prominent, we must also learn when they are useful to the user, and call attention to them only when they will prove productive.

CONCLUSION

Based on past research on social information seeking, we have certain intuitions about how users should behave around social annotations: they should find them broadly useful, and they should notice them. Our results indicate that, in reality, users behave in a more nuanced way.

Our first study yielded two unexpected results. First, in some contexts, social annotations shown on search result pages can be useless to searchers. They disregard information from people who are strangers, or unfamiliar friends with uncertain expertise. Searchers are looking for opinions and reviews from knowledgeable friends, or signs of interest from close friends on hobbies or other topics they have in common.

The more counterintuitive result from our first study was that subjects did not notice social annotations. From our second experiment, we were able to conclude that this unawareness

was mainly due to specialized attention patterns that users exhibit while processing search pages. Users deconstruct the search results: they pay attention to titles and URLs and then turn toward snippets and annotations for further evidence of a good result to click on. Moreover, the reading of snippets and annotations appears to follow a traditional top-to-bottom reading order, and friend pictures that are too small simply blend into snippets and become part of them. These focused attention behaviors seem to derive from the task-oriented mind-set of users during search, and might be explained by the effect of *inattentional blindness* [28]. All of this makes existing social annotations slip by, unnoticed.

IMPLICATIONS AND FUTURE DIRECTIONS

Our findings have implications for both the content and presentation of social annotations.

For content, three things are clear: not all friends are equal, not all topics benefit from the inclusion of social annotation, and users prefer different types of information from different people. For presentation, it seems that learned result-reading habits may cause blindness to social annotations. The obvious implication is that we need to adapt the content and presentation of social annotations to the specialized environment of web search.

The first adaptation could target broad search-topic categories: social annotations are useful on easily-identified topics such as restaurants, shopping, local services, and travel. For these categories, social annotations could be made more visually prominent, and expanded with details such as comments or ratings.

Our observation that the friend's topical expertise affects the user's perception of the social annotation and search result relevance allows for additional, fine-grained adjustments. With lists of topics on which friends are knowledgeable, we could give their annotations more prominence on those topics.

The areas of expertise or interest of a specific user could either be provided explicitly by the user (typically not during web search, but other interactions such as sign-up flows) or inferred implicitly from content created or frequently consumed by the user (e.g., authored posts on social networks, news feed subscriptions, exchanged emails, visited web pages). The inference can be done using standard text classification or clustering techniques [29].

To achieve the desired effect, we can manipulate the presentation of social annotations in a variety of ways to give them more prominence. For instance, we can increase the picture size, change the placement of the annotation within the search result, or alter its wording and information content.

In the future, we would like to conduct a third experiment to test this newly-gained understanding of social annotations. Using the insights from the second experiment, we could design a study in which social annotations are prominent. Then, we could test the qualitative claims of our first experiment by showing annotations from different types of contacts, on different verticals and topics.

A brave new world of social search is upon us. As the web becomes more social, social signals will factor into an ever-increasing part of our search experience. It is our hope that the knowledge obtained in these two studies will push the frontiers of social annotations in web search forward.

REFERENCES

1. Ackerman, M., and Malone, T. Answer garden: A tool for growing organizational memory. In *Proc. ACM SIGOIS and IEEE CS TC-OA* (1990), 3139.

2. Ackerman, M. S., and McDonald, D. W. Answer garden 2: merging organizational memory with collaborative help. In *Proc. CSCW*, CSCW '96, ACM (New York, NY, USA, 1996), 97105. ACM ID: 240203.

3. Amershi, S., and Morris, M. R. CoSearch. In *Proc. SIGCHI*, ACM Press (2008), 1647.

4. Bao, S., Xue, G., Wu, X., Yu, Y., Fei, B., and Su, Z. Optimizing web search using social annotations. In *Proc. WWW* (2007), 501510.

5. Borgatti, S., and Cross, R. A relational view of information seeking and learning in social networks. *Management science* (2003), 432445.

6. Carmel, D., Zwerdling, N., Guy, I., Ofek-Koifman, S., Har'el, N., Ronen, I., Uziel, E., Yogev, S., and Chernov, S. Personalized social search based on the user's social network. In *Proc. CIKM* (2009), 12271236.

7. Chi, E., Gumbrecht, M., and Hong, L. Visual foraging of highlighted text: An eye-tracking study. In *Proc. HCI* (2007), 589598.

8. Chi, E. H. Information seeking can be social. *Computer 42*, 3 (2009), 4246.

9. Cutrell, E., and Guan, Z. What are you looking for? In *Proc. SIGCHI*, ACM Press (2007), 407.

10. Erickson, T., and Kellogg, W. A. Social translucence: an approach to designing systems that support social processes. *ACM ToCHI 7*, 1 (Mar. 2000), 59–83.

11. Erickson, T., Smith, D. N., Kellogg, W. A., Laff, M., Richards, J. T., and Bradner, E. Socially translucent systems. In *Proc. SIGCHI*, ACM Press (1999), 72–79.

12. Evans, B. M., and Chi, E. H. Towards a model of understanding social search. In *Proc. CSCW* (2008), 485494.

13. Evans, B. M., Kairam, S., and Pirolli, P. Do your friends make you smarter?: An analysis of social strategies in online information seeking. *Information Processing & Management 46*, 6 (Nov. 2010), 679–692.

14. Golovchinsky, G., Pickens, J., and Back, M. A taxonomy of collaboration in online information seeking. In *Proc. 1st Workshop on Collaborative Information Seeking* (2008).

15. Golovchinsky, G., Qvarfordt, P., and Pickens, J. Collaborative information seeking. *Information Seeking Support Systems* (2008).

16. Granka, L. A., Joachims, T., and Gay, G. Eye-tracking analysis of user behavior in WWW search. In *Proc. SIGIR*, SIGIR '04, ACM (New York, NY, USA, 2004), 478479.

17. Guan, Z., and Cutrell, E. An eye tracking study of the effect of target rank on web search. In *Proc. SIGCHI*, ACM Press (2007), 417.

18. Haines, R. A breakdown in simultaneous information processing. *Presbyopia research. From molecular biology to visual adaptation* (1991), 171175.

19. Hershler, O., and Hochstein, S. At first sight: A high-level pop out effect for faces. *Vision Research 45*, 13 (2005), 17071724.

20. Heymann, P., Koutrika, G., and Garcia-Molina, H. Can social bookmarking improve web search? In *Proc. intl. conf. on Web search and web data mining*, ACM Press (2008), 195.

21. Hill, W. C., Hollan, J. D., Wroblewski, D., and McCandless, T. Edit wear and read wear. In *Proc. SIGCHI*, ACM (Monterey, California, United States, 1992), 3–9.

22. Horowitz, D., and Kamvar, S. D. The anatomy of a large-scale social search engine. In *Proc. WWW*, Proc. WWW, ACM (New York, NY, USA, 2010), 431440. ACM ID: 1772735.

23. Joachims, T., Granka, L., Pan, B., Hembrooke, H., and Gay, G. Accurately interpreting clickthrough data as implicit feedback. In *Proc. SIGIR*, SIGIR '05, ACM (New York, NY, USA, 2005), 154161.

24. Kammerer, Y., Nairn, R., Pirolli, P., and Chi, E. H. Signpost from the masses: learning effects in an exploratory social tag search browser. In *Proc. SIGCHI* (2009), 625634.

25. Langton, S., Law, A., Burton, A., and Schweinberger, S. Attention capture by faces. *Cognition 107*, 1 (2008), 330342.

26. Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L., and Gay, G. The influence of task and gender on search and evaluation behavior using google. *Information Processing & Management 42*, 4 (July 2006), 1123–1131.

27. Mack, A., Pappas, Z., Silverman, M., and Gay, R. What we see: Inattention and the capture of attention by meaning. *Consciousness and Cognition 11*, 4 (2002), 488506.

28. Mack, A., and Rock, I. Inattentional blindness. In *Visual Attention*. The MIT Press, 1998.

29. Manning, C. D., Raghavan, P., and Schutze, H. *Introduction to information retrieval*. Cambridge University Press, 2008.

30. Morris, M. R. A survey of collaborative web search practices. In *Proc. SIGCHI*, ACM Press (2008), 1657.

31. Morris, M. R., and Horvitz, E. SearchTogether. In *Proc. UIST*, ACM Press (2007), 3.

32. Morris, M. R., Lombardo, J., and Wigdor, D. WeSearch. In *Proc. CSCW*, ACM Press (2010), 401.

33. Nelson, L., Held, C., Pirolli, P., Hong, L., Schiano, D., and Chi, E. H. With a little help from my friends: examining the impact of social annotations in sensemaking tasks. In *Proc. SIGCHI*, CHI '09, ACM (New York, NY, USA, 2009), 17951798. ACM ID: 1518977.

34. Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., and Granka, L. In google we trust: Users decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication 12*, 3 (2007), 801823.

35. Pickens, J., Golovchinsky, G., Shah, C., Qvarfordt, P., and Back, M. Algorithmic mediation for collaborative exploratory search. In *Proc. ACM SIGIR*, ACM Press (2008), 315.

36. Pirolli, P. An elementary social information foraging model. In *Proc. SIGCHI* (2009), 605614.

37. Pirolli, P., and Card, S. Information foraging. *Psychological Review 106* (1999), 643–675.

38. Rock, I., Linnett, C. M., Grant, P., and Mack, A. Perception without attention: Results of a new method. *Cognitive Psychology 24*, 4 (Oct. 1992), 502–534.

39. Silvers, V., and Kreiner, D. The effects of Pre-Existing inappropriate highlighting on reading comprehension. *Reading Research and Instruction 36*, 3 (1997), 21723.

40. Simons, D. J., and Chabris, C. F. Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception 28*, 9 (1999), 1059–1074. PMID: 10694957.

41. Theeuwes, J., and Van der Stigchel, S. Faces capture attention: Evidence from inhibition of return. *Visual Cognition 13*, 6 (2006), 657665.

42. Von Restorff, H. Ueber die wirkung von bereichsbildungen im spurenfeld. *Psychological Research 18*, 1 (1933), 299342.

43. Yanbe, Y., Jatowt, A., Nakamura, S., and Tanaka, K. Can social bookmarking enhance search in the web? In *Proc. ACM/IEEE-CS joint conf. on Digital libraries*, ACM Press (2007), 107.

44. Zanardi, V., and Capra, L. Social ranking: uncovering relevant content using tag-based recommender systems. In *Proc. ACM RecSys*, ACM Press (2008), 51.