All the News that's Fit to Read:
A Study of Social Annotations for News Reading.

Chinmay Kulkarni
Stanford University HCI Group
353 Serra Mall, Stanford, CA
chinmay@cs.stanford.edu

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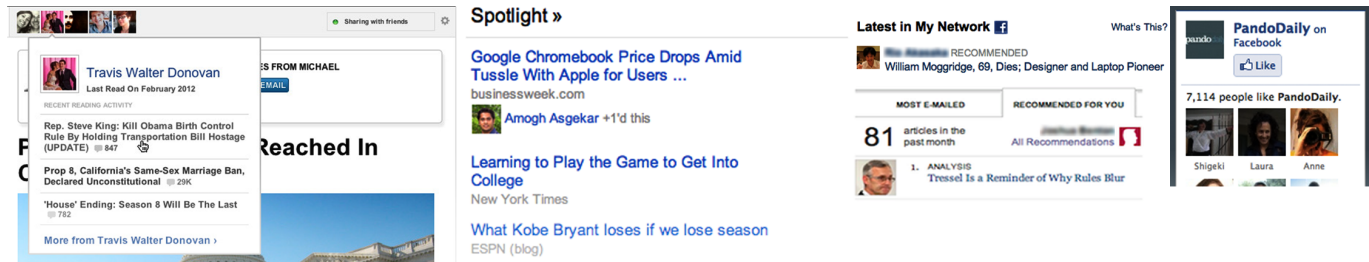Mountain View, CA
edchi@google.com

Figure 1. Despite the ubiquity of social annotations online, little is known about their effects on readers, and relative effectiveness. From left: (1) Facebook Social Reader showing articles friends recently read; (2) Google News Spotlight, algorithmic recommendations combined with annotations from friends; (3) New York Times recommendations for a logged-in user, from friends(top), and algorithms(bottom); (4) Facebook widget showing annotations from strangers for non-logged-in user.

ABSTRACT

As news reading becomes more social, how do different types of annotations affect people's selection of news articles? This paper reports on results from two experiments looking at social annotations in two different news reading contexts. The first experiment simulates a logged-out experience with annotations from strangers, a computer agent, and a branded company. Results indicate that, perhaps unsurprisingly, annotations by strangers have no persuasive effects. However, surprisingly, unknown branded companies still had a persuasive effect. The second experiment simulates a logged-in experience with annotations from friends, finding that friend annotations are both persuasive and improve user satisfaction over their article selections. In post-experiment interviews, we found that this increased satisfaction is due partly because of the context that annotations add. That is, friend annotations both help people decide what to read, and provide social context that improves engagement. Interviews also suggest subtle expertise effects. We discuss implications for design of social annotation systems and suggestions for future research.

Author Keywords

Social Computing; Social annotation; news reading; recommendations experiment; user study.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Newspaper websites like the New York Times allow readers to recommend news articles to each other. Restaurant review sites like Yelp present other diners' recommendations, and now several social networks have integrated social news readers. Just like any other activity on the Web, online news reading seems to be fast becoming a social experience.

Internet users today see recommendations from a variety of sources. These sources include computers and algorithms, companies that publish and aggregate content, their own friends and even complete strangers (See Figure 1 for a sampling of recommendations online). Annotations are endorsements by other agents (like other users and computers), and are increasingly used with content recommenders. Social annotations (i.e. endorsements by other users) have become especially popular with content recommenders using social signals [28, 29, 4].

In addition to providing recommendations, websites also often share their users' reading activity, and this too happens in a number of different ways. Users may need to share what they read explicitly (e.g. by clicking a 'Share' or 'Like' button), or websites may share such activity implicitly and automatically.

Given the ubiquity of online social annotations, it is surprising how little is known about how social annotations work for online news. While there have been some studies of annotations [21, 22], to our knowledge, there is no good published research on the engagement effects of annotations in social news reading. How do annotations help people decide what

articles they read? Does recording and sharing what someone reads with others affect their decisions of what they read? And do different types of annotations, such as those from algorithms and people affect readers differently?

Intuitively, different types of annotations offer different explanations to the user, and these explanations should have different persuasive effects. For example, algorithmic annotations may bestow a sense of impartiality [27], branded annotations (such as the New York Times) may bestow authority and reputation [24]. And annotations by other people may be persuasive due to social influence [14]. In short, the kind of annotation may affect people's reading behavior.

Moreover, in many social readers, the reading decisions are recorded (and shared) automatically. This may cause users to be more cautious. In a social context, Goffman's work suggests that users will think about how they appear to others [9]. This suggests that the presence of behavior recording may interact with annotation.

To understand the future of social news reading, we investigated how these different forms of annotations might affect user behaviors. From a system designer's point of view, it is important to support both users who are logged-in and those that are not, especially because users that aren't logged in may constitute the bulk of traffic.

System designers only have a limited set of annotation types for users who are not logged-in. While designers can show logged-in users *personalized* recommendations with annotations from their friends, they can only show users who aren't logged-in *non-personalized* recommendations with annotations that are branded, algorithmic or from strangers.

Therefore, we conducted two separate experiments, focusing on the non-personalized and the personalized experiences separately. We also varied the experimental condition so that, for half the participants, their reading actions were visibly recorded with a feedback message (e.g. "You read this (publicly recorded).").

In the first experiment, we first investigated the logged-out non-personalized experience. Importantly, in this experiment, we held the news stories constant, while only varying the annotations shown to the subjects. Participants saw annotations from strangers, a computer algorithm and a fictitious company.

We chose to show a fictitious (yet news-related) company in order to examine the general effects of using annotations by companies, rather than the effect of specific brands. We know from marketing studies that different companies have different brand perceptions, for instance, the New York Times, the Guardian, the Washington Post, Fox News, and the Onion represent very different brands of news, and users evaluate them differently. It is conceivable that effects may well depend on these perceptions, but it is outside of the scope of this paper to examine all of the branding effects, which are best studied in a marketing study.

Results from this experiment suggest that annotations by strangers, perhaps unsurprisingly, have no persuasive effects.

However, both the computer program and, surprisingly, the unknown branded company annotations had a persuasive effect. For the computer program, we surmise that it is viewed as being impartial and unbiased. However, the fictitious company having an effect was surprising because one might argue that it is really not that different from a total stranger (and potentially with a hidden agenda!)

In the second experiment, we simulated a logged-in personalized context by presenting users with real recommendations that were annotated by real friends. We found that social annotations by friends are not only persuasive but also improved user satisfaction. In the interviews, we found that this increased satisfaction is driven in part by the context that annotations add. We also find evidence for thresholding—social annotations have their persuasive effects when the expertise or tie strength of the annotator exceeds a threshold, but the precise identity of the annotator is not important.

RELATED WORK

Annotations as decision aids

Annotations can be seen as as decision aids that provide proximal cues to help people find distal content. Research on annotations as decision-tools focuses primarily on web-search, but some research exists on news-reading.

When a clear information need is present (e.g. with web-search), annotations are seen primarily as tools that help people decide which resources are suitable to their current information need [11, 21].

Prior work has focused on two aspects: (a) how annotations should be presented, and (b) which annotations are useful as decision aids. For presentation of annotations, the reading order of annotations determines when (and if) annotations are seen [21].

In the presence of a clear information need, social annotations are most helpful when they are from people known to have expertise in the current search domain (e.g. professional programmers for questions about programming, or "foodies" for restaurant recommendations) [21]. Such expertise plays a more important role than social proximity to annotators [15].

When information needs are not as specific, annotations may be processed differently (as with online news reading). For a news recommendation system, information needs are often vague and exploratory, both for experts and journalists [12, 7], and for news consumers [26]. Because the information need is vague, people use proximal cues to find resources that maximize parameters such as verity, importance, or interestingness. Prior work has identified that journalists use cues such as social proximity and geographic location to estimate the verity of social-media news [7], and readers use explicit popularity indicators to decide if news is interesting [14]. Our research adds to this knowledge by studying the effects of annotations in their role as proximal cues.

In addition, online news websites often have annotations by agents other than other users friends (such as by news companies, editors, etc). Sundar et. al. show that the news source
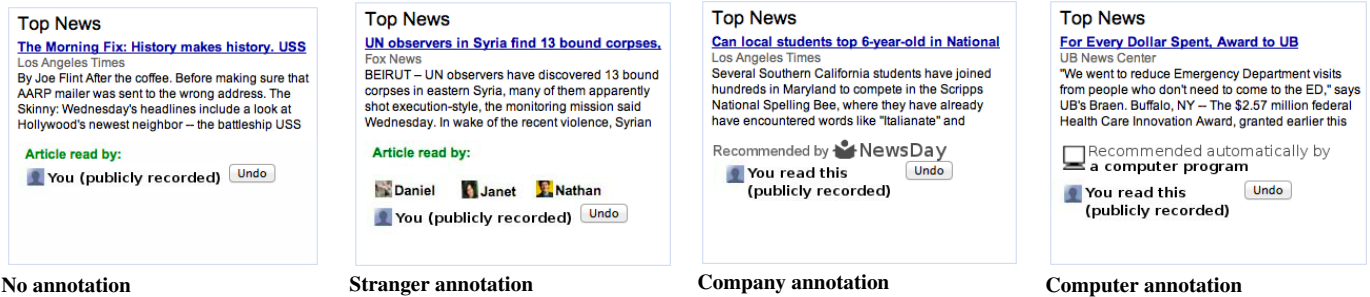
Figure 2. News articles in different annotation conditions, shown with recording present. The "You read this (publicly recorded)" notice appears when users click the headline. The no-recording conditions were the same except that they did not show the notice when users clicked on the headline.

affects how believable readers think the news is, how interesting they found it, etc. *after* they consume the news [27]. Somewhat surprisingly, we didn't find any published work on how these different sources act as proximal cues. Our current experiment adds to this theory, by examining how annotations affect readers *before* they read the news, and in particular how they affect reader's reading decisions. Our experiments also inform theory and design about engagement effects of *social* annotations for news, a topic which has been ignored in prior work.

Annotations as persuasion

Annotations can be also be seen as a way to *persuade* people to take certain actions. This view is motivated by prior work that demonstrates that people alter their choices [30], change their reported ratings in the face of opposing social opinion [6], or even engage in entirely new activities [20]. This prior work suggests ways to make social annotations more persuasive. For instance, Golbeck et al. demonstrate that displays of social information, especially expertise clues, help build trust and persuasiveness [10].

While this past research equips practitioners with advice about how to display annotations to end-users to improve perceived trust, it doesn't provide guidance about which annotations to show. In addition, prior research ignores many of the annotation types that are now prevalent, such as those by brands or computers. While past research suggests annotations maybe persuasive [24, 8], their relative trade-offs have not received attention.

The constrained view of annotations as persuasion alone also ignores other benefits that annotations may have. Therefore, naively maximizing persuasiveness may reduce annotations' helpfulness at identifying good content (by making both good and bad content seem equally persuasive).

Annotations as Social Presentation

Goffman notes that humans change their behavior in social situations to reflect how they want to be seen [9]. Such social presentation has also been observed in online social networks [18, 16]. Since social annotations are endorsements of content, they may be seen as a form of social presentation by people who annotate content.

Social presentation may be equally important for readers of annotated content. When systems share reading behavior with other people, users may engage in privacy regulation [3, 23], which may change their reading behavior according to content and the audience with which such behavior is shared.

We aim to build on past work to increase our knowledge about how reading behavior changes when it is recorded and shared publicly. This paper reports on studies of readers in two different situations in two experiments.

EXPERIMENT 1: NON-PERSONALIZED ANNOTATIONS

Goals
Our first experiment studies how people use annotations when the content they see is not personalized, and the annotations are not from people in their social network. This is the case when users see annotated content when they are not logged-in to a social network.

Participants
We performed this experiment on Amazon's Mechanical Turk platform. We selected participants who were US-resident workers on Mechanical Turk. A total of $N = 560$ participants (237 female) took part in the experiment, and were compensated US$0.50 each for their participation. As a prerequisite, we asked participants to confirm they could communicate in English. The experimental platform captured participants Amazon Worker IDs to ensure that participants could not participate in the experiment more than once.

Procedure

Conditions
The experiment manipulated two variables: Annotation Type (four levels: None, Computers, Strangers, Company) and Recording (Present or Absent) in a 4X2 Between-subjects design. We used a between-subjects design because, on Mechanical Turk, it is difficult to ensure that participants complete all conditions in a within-subjects experiment.

Setup and procedure
At the start of the experiment, participants were told we were testing an experimental news system, which would show them different news articles. Participants saw four pages of news headlines, with six boxes of news articles on each page. Each article box was annotated based on the annotation condition the participant was in (Figure 2 shows the boxes for all annotation conditions where recording was present).
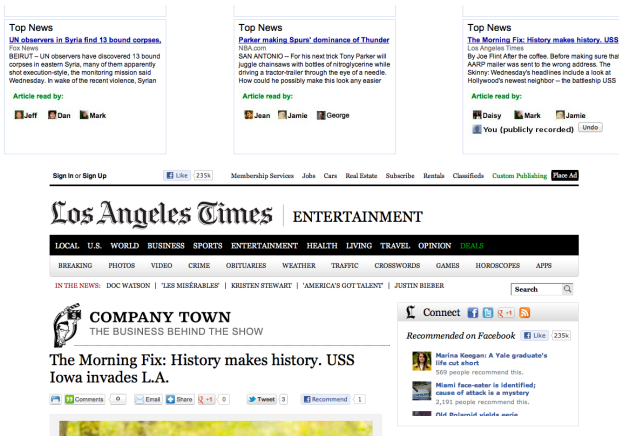
Figure 3. Experimental setup: Clicking on a headline opened the linked article in a frame (screen-shot shows an article from Los Angeles Times).
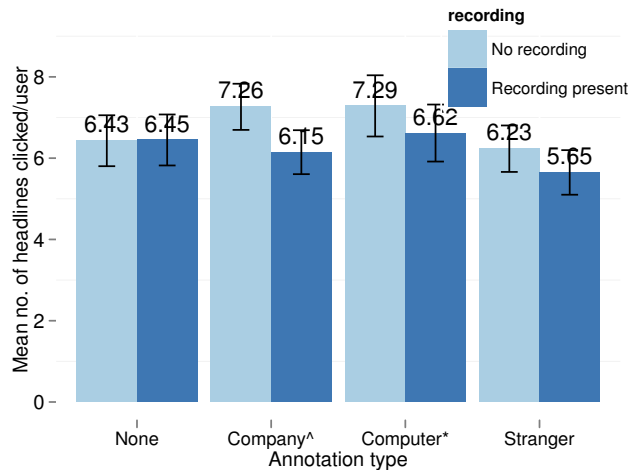

Figure 4. Participants clicked significantly more headlines when annotated by a Computer, and marginally more with a fictitious company, but annotations by strangers had no effect. Recording reduces the number of clicks when annotations are present.

The company annotation used a company name that was chosen to be not familiar to participants, and yet could be a plausible company that made news recommendations.

Participants were told to click on the articles that they thought were interesting. When participants clicked a box, the selected article opened in a browser frame, below all the other boxes (See Figure 3). Clicked boxes also had an *Undo* button if they clicked on an article-box in error.

Participants in conditions where Recording was Present were told that others in the experiment would see their name displayed next to articles they read. Upon clicking an article box, they saw a recording indicator in the article box: "You read this (publicly recorded)".

Unknown to the participants, participants across all conditions saw the same set of news articles. These news articles were taken from the day's headlines (from Google News) in six different categories: Health, National, World, Entertainment, Technology, and Sports. Each page displayed one news item from each category.

Results
To analyze the results of this experiment, we compared the number of articles participants clicked across different conditions. In our experimental setup, participants were not required to click a minimum number of articles, and there were 252 participants who clicked no articles. The numbers of such participants was independent of condition ($t(7) = 0, p < 1$). To get reliable results, the analysis below discards such participants (list-wise deletion [2]). The number of articles participants clicked differed between the annotation conditions ($X^2(3, N = 308) = 12.89, p < 0.05$, Type II ANOVA [17])

Computer, company annotations increase articles clicked
Surprisingly, annotation by Computers increased the number of articles clicked ($t_{300} = +2.03, p < 0.05$) compared to the None condition, and Company annotations had a similar, marginally significant effect ($t_{300} = +1.93, p = 0.05$). Participants clicked a mean of 6.98 and 6.73 articles in the computer, and company condition, respectively compared with

6.43 articles in the None condition (SD=4.25, 3.36, and 3.96 respectively).

Strangers' annotations don't affect number of articles clicked
Annotations by Strangers had no effect on the number of headlines participants clicked ($t_{300} = -0.49, p > 0.6$).

Recording reduces the number of articles clicked whenever annotations are present
While Recording had no main effect on the number of articles read, it had significant interaction effect on the number of articles clicked when Company annotations were present, compared to the None condition ($t_{300} = -2.46, p < 0.05$). Participants clicked a mean of 7.26 articles with no Recording, vs. 6.15 when Recording was present. Recording also reduced the number of headlines clicked for Computer annotations, but the effects were not significant.

Discussion
Clicks as measure of persuasiveness
All participants saw the same articles, and these articles were displayed identically except for Annotations and Recording conditions, with the same number of words in title and snippets. Prior work has shown that relative click-rates are an accurate measure for user-intent [13, 19]. Therefore, we use number of articles clicked in each condition as a measure of how persuasive an annotation is in making people read articles (in the presence and absence of Recording).

The presence of Recording reduces the persuasiveness of annotations, which suggests that news reading is considered a social activity where participants engage in social presentation [9]. We were surprised then that annotations by unknown companies and computers were persuasive, but those by unknown people (strangers) weren't, because the theory does not predict this difference [25].
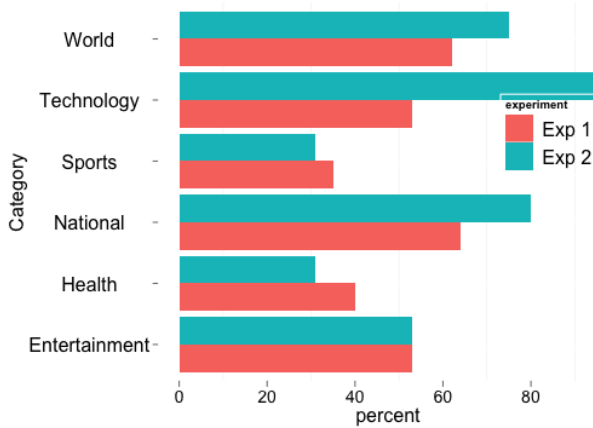
Figure 5. General news reading behavior reported by participants in both experiments (numbers add to more than 100% since participants could select more than one category)

DO FRIENDS AND PERSONALIZATION MATTER?

Overall, Experiment 1 led to the somewhat surprising result that while annotations by companies and computers are persuasive, those by strangers are not. From a practical point of view then, annotations by computers and companies may be more valuable in a logged-out context.

For a user that is logged-in, system designers can provide recommendations that are personalized based on social signals and annotated by real friends. Our second experiment studies such a logged-in, personalized environment.

EXPERIMENT 2: PERSONALIZED ANNOTATIONS

Goals

Our second experiment studies how people use annotations in personalized contexts with annotations from friends. In particular, it asks two questions. First, as decision aids, do personalized social annotations help people discover and select more interesting content? Second, as persuasion, are annotations by friends persuasive, even though our first experiment suggests those by strangers are not?

Participants

We recruited participants from amongst employees at our organization. All participants lived in the US, spoke English, and worked in non-technical positions, such as managers, receptionists, support staff. For participation in our experiment, we raffled three $50 gift cards for Amazon as compensation.

While this second experiment was run on a separate participant pool, the two are largely similar in geographical distribution (US resident), gender (42% female in Experiment 1, and 39% female in Experiment 2) and age (median 30 in Experiment 1, and 28 in Experiment 2). Participants in both experiments reported they read the same kinds of news articles, except for Technology, which was read more frequently by participants in Exp 2 (Figure 5). Participants in Experiment 2 reported they forwarded/shared content more often (29% shared "at least once a day", vs. 18.2% for Experiment 1). Participants in both experiments were equally likely to

share content with "close friends and family" (85% in Exp 1, 72% in Exp 2), but participants in Exp 2 were more likely to share with Coworkers (21% in Experiment 1, vs. 58% in Experiment 2.) Both sets of users reported they received an interesting article from people in their social network equally frequently (the median choice was "Few times every week", and was chosen by 25% in Experiment 1, and 30% in Experiment 2.)

Because Experiment 2 relies on content that is drawn from a participant's own social network, we used a different participant pool. However, both participant pools are similar enough in news *consumption* behavior that both experiments can together contribute to our understanding of user behavior (Using a local pool also allowed us to interview participants).

We wanted to show participants a list of news articles that was actually personalized to them, and annotations from real friends. Therefore, we initially solicited a much larger number of employees to allow us to look at public +1's by their contacts on Google+. Then, amongst those that responded to our call, we selected participants where we could find at least 12 news-like URLs that their friends had shared. This filtering process resulted in $N = 59$ participants.

Procedure

The procedure for this experiment was largely similar to that of Experiment 1, i.e. participants were told our research team was evaluating a news recommendation system, and they should click articles that interested them. We highlight the major differences from Experiment 1 below.

Conditions and measures

This experiment used a mixed between- and within-subjects manipulation. We manipulated three independent variables: Annotation Types (None, Friend or Stranger), whether Recording was present or absent, and whether news stories chosen were Personalized or not. Recording was a between-subjects variable, while the other two variables were manipulated as within-subjects variable.

We decided to only include certain combinations of variables (Table 1). This was done partially because Non-personalized X Friend combination might be strange to users when they see annotations on articles that did not make sense for their friend.

Within-subjects	Between-subjects
Personalized X $< none, friend >$ Non-personalized X $< none, stranger >$	Recording (Present/Absent)

Table 1. Conditions in Experiment 2.

This experiment used two dependent measures for engagement: the number of articles participants clicked, and a rating of interestingness for each article. After the participant clicked on an article, they were asked to rate the article on a Likert scale of 1-5 (with 5 being extremely interesting). While rating was optional, all clicks we captured had associated interestingness ratings.

Setup and procedure
Participants saw 24 articles on four pages as in Experiment 1. However, each page showed articles from a different within-subjects condition (with the condition order counterbalanced). Because Recording was a between-subjects variable, participants either saw all articles with recording or without it. All article-boxes used the same length for the headline and news snippet.

When participants clicked an article box, the article opened in a frame below (similar to Figure 3). Upon clicking on an article, we showed an *Undo* button (similar to Experiment 1), and a widget to rate the interestingness of the article on a Likert scale.

Pages with personalized content (Personalized X Friend annotation, and Personalized X None) showed articles that were +1'd by the participant's friends. The Friend annotation showed the name of the friend who had publicly +1'd the article. Pages without personalized content (Non-personalized X Stranger annotation, and Non-personalized X None) showed articles from Google News, similar to Experiment 1. Each page showed one article from each of the six categories.

Hypotheses
Since this experiment didn't use a fully-crossed design, we analyzed data with planned comparisons. All analyses used a mixed effects model for repeated measures, with participants having a fixed intercept. We had six hypotheses for our planned comparisons.

H1: Personalization If we don't show annotations, personalization (based on social signals) increases engagement
H2: Friend Annotations If we only show personalized articles, showing the friends' annotation increases engagement
H3: Stranger Annotations If we only show non-personalized articles, showing strangers' annotations increases engagement
H4: Net effect Personalization and friend annotations together increase engagement over non-personalized, unannotated content
H5: Recording Recording reduces people's engagement levels
H6: Recording interaction Recording interacts with annotation or personalization

We consider these comparisons to be simultaneous, and so use the Holm-Bonferroni correction to control the family-wise error rate [1].

Results
H1: Personalization
Without annotations, Personalization had no significant effect on the number of articles clicked. Without annotations, both kinds of content, personalized and non-personalized received about the same number of clicks, $(t(58) = 0.87, p > 0.2)$. Participants clicked a mean of 1.91 non-personalized articles, and 1.98 personalized (SD=1.45, 1.43 respectively).

Similarly, personalization had no significant effect on interestingness $(t(138) = -0.36, p > 0.5)$. Both kinds of content were rated similarly without annotations: mean=3.17 for non-personalized, 3.11 for personalized (SD=1.09 and 1.19 respectively).

Therefore, H1 is not supported for clicks or interestingness, so personalized content (based on social signals) doesn't change engagement, surprisingly.

H2: Friend Annotations
With Personalization, subjects in the Friend annotations condition clicked on more articles $(t(58) = 1.62, p = 0.05$, which is above the Holm-Bonferroni correction threshold for significance). Participants clicked a mean of 1.98 for non-annotated articles, and 2.16 articles for Friend annotated articles (SD=1.43 and 1.49 respectively).

Participants rate articles that are presented with Friend annotations as more interesting than those without annotation $(t(138) = 3.96, p < 0.01)$ [mean=3.11 for non-annotated, 3.61 for friend (SD=1.19 and 1.11 respectively)].

Therefore, H2 is marginally supported for clicks, and supported for interestingness. That is, showing Friend annotations appears to increase user engagement for Personalized content.

H3: Stranger Annotations
Stranger annotations made people click marginally more articles $(t(58) = 1.63, p = 0.05$, which is above the Holm-Bonferroni correction threshold for significance). [mean=1.98 for non-annotated articles, and mean=2.16 for articles with stranger annotations (SD=1.43 and 1.48 respectively).]

Participants rate articles annotated by strangers *lower* than articles without annotations $t(138) = -2.46, p < 0.01)$ [mean=3.17 for non-annotated, 2.79 for stranger (SD=1.09 and 1.07 respectively)].

Therefore, H3 is marginally supported for clicks, and not supported for interestingness. Showing strangers' annotation increased click through, but ultimately decreases interestingness.

H4: Net effect
Participants clicked on personalized content with friend annotations more than non-personalized with no annotation $(t(58) = 2.135, p < 0.05)$. Participants clicked a mean of 1.91 articles that were non-annotated and non-personalized vs. 2.16 articles that were personalized with friend annotations (SD=1.45 and 1.50 respectively).

Similarly, participants rated personalized content with friend annotations to be more interesting than non-personalized with no annotation $(t(138) = 3.67, p < 0.01)$. On average, non-annotated, non-personalized articles had a rating of 3.17 vs. a mean rating of 3.67 for personalized content with friend annotations (SD=1.09 and 1.11 respectively).

Therefore, H4 is supported for clicks and interestingness, suggesting that personalization and annotation together work hand-in-hand to provide for better user experience overall.

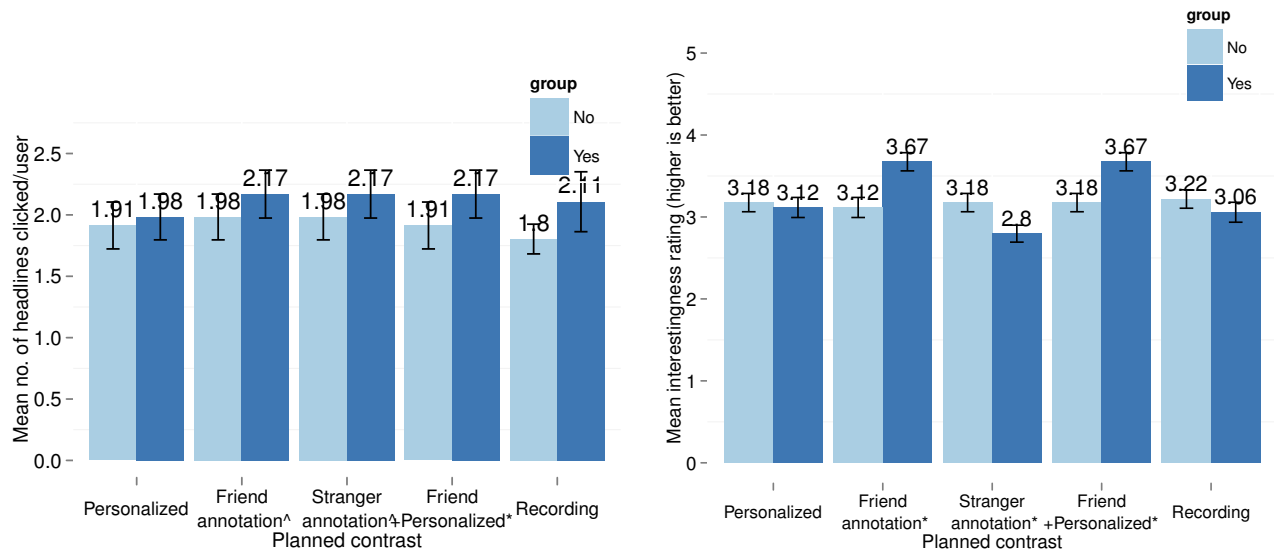*H5: Recording and H6: Recording Interaction*

Figure 6. Friend and stranger annotations marginally increase the number of articles users click. Friend annotations increase the rated interestingness of articles, while stranger annotations decrease it.

With the Holm-Bonferroni correction, we found no significant main or interaction effects of recording, for either clicks or interestingness.

Therefore, H5 and H6 are not supported, suggesting recording doesn't significantly affect engagement, in surprising contradiction with Experiment 1.

Summary of findings
Results from this experiment suggest that friend annotations help engagement while those by strangers, while marginally increasing the number of articles clicked, do not improve ratings for interestingness. It also presents evidence that suggests that personalization and annotation work together to improve user experience overall.

To augment these findings with qualitative descriptions of how annotations are used, we conducted interviews with a subset of participants.

INTERVIEWS
For our interviews, we chose eight participants from Experiment 2 at random (4 female). All interviews were conducted by phone, and took approximately 20 minutes each. Participants were not compensated separately for interviews. All interviews took place within two hours of participants completing the experiment.

Interviews used a retrospective-think-aloud (RTA) with critical-incident method. Among the list of articles clicked by the participant, the researcher picked an article at random from each of the conditions the participant was exposed to. Participants were then asked to describe the article, and why they found it interesting. During the interview, the researcher asked probing questions, like "why did you click the article?", "did you notice the annotation?" or "who was the annotator?" In some cases, the participant mentioned annotations without such questions.

Based on these interviews, we surmise that annotations made participants read articles primarily in three cases. First, when the annotator was above a threshold of social closeness; second, when the annotator had subject expertise related to the news article; and third, when the annotation provided additional context. We describe each of these below.

Annotators above a threshold of social proximity
Interviewees frequently remembered that a given article was annotated by a friend, but did not recollect the identity of the annotator. This suggests that while annotations by friends are useful, they are used more as a thresholding filter.

We found one exception to this pattern: Participants remembered annotators that were close friends. In addition, interviewees often said they were willing to "take a chance" on such annotators. For instance, one participant said, "I never watch videos…but I'll read most things Krystal recommends."

Annotators with subject expertise
Similar to prior work on social annotations in web search, we find that participants read content that was annotated by social contacts with expertise [21, 15]. 3 of 8 participants reported they clicked on an article because that was annotated by a friend who had expertise in the area. Unlike search, however, this expertise was related to the article the annotation was on, rather than the user's information need. In fact, participants sometimes read articles they otherwise would not, because they were annotated by a subject expert. For example, one participant said "Doug is a friend of mine, and is a cartoonist. If Doug is reading that cartoon, then I'm going to…"

Annotations that provided context
Annotations also add context to recommendations. For instance, one participant revealed he read an article about a new railroad in Philadelphia because his "friend from Philly"

had annotated it. Similarly, office conversations about standing desks made another participant click an article about their health benefits which was annotated by her colleague.

DISCUSSION

Our two experiments extend our knowledge about social annotations. They show that people's reading behavior are affected not only by the way recommendations are generated (e.g. recommendations from friends or top headlines), but also by designers' choices about how these annotations are displayed. While the effects that recommendations and annotations have somewhat overlapping effects, below we elaborate on three particular roles social annotations play.

Annotations as persuasion

Results from Experiment 1 suggest that in a logged-out context, annotations by strangers don't persuade people to click. Surprisingly, those by computers and even companies were persuasive.

In our second experiment, we see that Strangers' and friends' annotations both marginally encourage people to click (in contrast to Experiment 1). Why the difference? One possibility is that participants in Experiment 1 (and in general in a logged-out context) know that other people are strangers. In contrast, in a logged-in context, participants may find it difficult to distinguish between strangers and distant acquaintances. This ambiguity may also have been exacerbated because the experiment design was within-subjects, and participants saw both friends and strangers.

Annotations for engagement

While both stranger and friend annotations marginally increase click rates, participants' rating for interestingness are a different story. While annotations from friends increase interestingesss, those by strangers decrease it. One possible explanation is that, even though strangers and friends have similar social proof effects (and so, persuasiveness), because strangers lack homophily, so their annotations do not increase the user's enjoyment. Another is that annotations work as descriptive social norms. Such norms involve perceptions not of what others approve but of what others actually do, and are also known to influence compliance decisions powerfully [5].

This homophily may also lead to annotations by friends providing additional context (as reported by our participants). In contrast, annotations by strangers fail to provide context, and may lead to people feeling cheated or confused about why content was annotated.

In our experiment, Personalized content didn't affect engagement by itself, but this could be the specific implementation of personalization we used, or because of domain effects. That is, we speculate that, in the news domain, we are less likely to elicit hate or love reactions as strongly as film and music.

Annotations as social presentation

In Experiment 1, we found that Recording, by itself, had no significant effects. However, Recording interacted with annotations that were persuasive. On the other hand, our results in Experiment 2 were not conclusive, but suggest that recording might reduce clicks. This conflicts with results from Experiment 1, and deserves further study.

One possibility is motivated by Altman's theory [3]: privacy regulation is a dynamic process that depends on circumstances. Therefore, participants in our second experiment, being employees of the same organization, may have implicitly trusted the experimental platform more. Further study of this phenomenon is important both because it has privacy implications and it may help designers create systems where people may share more openly.

Annotations and recommendations

This paper also shows that people's reading behavior is affected both by the way recommendations are generated (e.g. recommendations from friends or top headlines), and by designers' choices about how these recommendations are displayed.

Limitations

Our experiments are a first examination of the role of annotations in news reading, and were therefore designed to identify high-level effects. Further investigation may highlight more nuanced effects. For instance, not all Google+ friends are alike. While our interviews provide some clues about the role of social proximity, future experiments could better quantify this role. Similarly, while our choice to show a fictitious, news-related company for company annotations demonstrates the effects of annotations by a topically-related (or even topic-expert) company, it may not generalize to companies that are not topically related.

CONCLUSION

Taken together, our experiments suggest that social annotations, which have so far been considered as a generic homogeneous tool to increase user engagement, are not homogeneous at all. Social annotations vary in their degree of persuasiveness, and their ability to change user engagement.

In a logged-out context, annotations by computers and companies are more persuasive than those by strangers. In a logged-in context, friend annotations are persuasive. Our interviews suggest that the most effective friend annotations are from those who are above a social proximity threshold, or by subject expert, or those that provide context.

Moreover, annotations go beyond persuasion and decision-making: they can make (social) content more interesting by their presence, at least in part by providing additional context to the annotated content.

This paper offers a first examination of the role of social annotations for news reading. Some questions for future research are: Does highlighting expertise help? Can the threshold for social proximity be algorithmically determined? If stranger annotations work because of social proof, does aggregating annotations (e.g. "110 people liked this") help? In addition, while this paper makes a first study of effects of annotations by company and computer algorithms, further research might reveal more nuances based on names of the companies and the presentation of these annotations.

There are several benefits of this and future research: not only will further understanding help designers create social recommender systems that are more enjoyable, it will also help us develop a theoretical background on how people make decisions about the way they stay aware of news in the presence of social information.

REFERENCES

1. Abdi, H. Holm's sequential bonferroni procedure. *Encyclopedia of research design (N. Salkind, DM Dougherty, and B. Frey, eds.). Sage, Thousand Oaks, California* (2010), 573–577.

2. Allison, P. D. *Missing Data*. Sage, Thousand Oaks, CA, 2001.

3. Altman, I., Vinsel, A., and Brown, B. Dialectic conceptions in social psychology: An application to social penetration and privacy regulation. *Advances in experimental social psychology 14* (1981), 107–160.

4. Bao, S., Xue, G., Wu, X., Yu, Y., Fei, B., and Su, Z. Optimizing web search using social annotations. In *Proceedings of the 16th international conference on World Wide Web*, ACM (2007), 501–510.

5. Cialdini, R. Descriptive social norms as underappreciated sources of social control. *Psychometrika 72*, 2 (2007), 263–268.

6. Cosley, D., Lam, S., Albert, I., Konstan, J., and Riedl, J. Is seeing believing?: how recommender system interfaces affect users' opinions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2003), 585–592.

7. Diakopoulos, N., De Choudhury, M., and Naaman, M. Finding and assessing social media information sources in the context of journalism. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, ACM (2012), 2451–2460.

8. Fogg, B. Persuasive technology: using computers to change what we think and do. *Ubiquity 2002*, December (2002), 5.

9. Goffman, E. The presentation of self in everyday life. 1959. *Garden City, NY* (1959).

10. Golbeck, J., and Fleischmann, K. Trust in social q&a: the impact of text and photo cues of expertise. *Proceedings of the American Society for Information Science and Technology 47*, 1 (2010), 1–10.

11. Google. Introducing google social search: I finally found my friend's new york blog! `http://googleblog.blogspot.com/2009/10/introducing-google-social-search-i.html`, 2009.

12. Hermida, A. Twittering the news: The emergence of ambient journalism. *Journalism Practice 4*, 3 (2010), 297–308.

13. Joachims, T., Granka, L., Pan, B., Hembrooke, H., Radlinski, F., and Gay, G. Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Transactions on Information systems 25*, 2 (2007).

14. Knobloch-Westerwick, S., Sharma, N., Hansen, D., and Alter, S. Impact of popularity indications on readers' selective exposure to online news. *Journal of Broadcasting & Electronic Media 49*, 3 (2005), 296–313.

15. Komanduri, S., Fang, L., Huffaker, D., and Staddon, J. Around the water cooler: Shared discussion topics and contact closeness in social search. In *Sixth International AAAI Conference on Weblogs and Social Media* (2012).

16. Lampe, C., Ellison, N., and Steinfield, C. A familiar face (book): profile elements as signals in an online social network. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2007), 435–444.

17. Langsrud, Ø. Anova for unbalanced data: Use type ii instead of type iii sums of squares. *Statistics and Computing 13*, 2 (2003), 163–167.

18. Lee, A., and Bruckman, A. Judging you by the company you keep: dating on social networking sites. In *Proceedings of the 2007 international ACM conference on Supporting group work*, ACM (2007), 371–378.

19. Li, L., Chu, W., Langford, J., and Schapire, R. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, ACM (2010), 661–670.

20. Masli, M., and Terveen, L. Evaluating compliance-without-pressure techniques for increasing participation in online communities. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, ACM (2012), 2915–2924.

21. Muralidharan, A., Gyongyi, Z., and Chi, E. H. Social annotations in web search. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems (CHI '12)* (New York, NY, 2012), 1085–1094.

22. Nelson, L., Held, C., Pirolli, P., Hong, L., Schiano, D., and Chi, E. With a little help from my friends: examining the impact of social annotations in sensemaking tasks. In *Proceedings of the 27th international conference on Human factors in computing systems*, ACM (2009), 1795–1798.

23. Palen, L., and Dourish, P. Unpacking privacy for a networked world. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2003), 129–136.

24. Raghubir, P., and Srivastava, J. Monopoly money: The effect of payment coupling and form on spending behavior. *Journal of Experimental Psychology: Applied 14*, 3 (2008), 213.

25. Reeves, B., and Nass, C. *The media equation: How people treat computers, television, and new media like real people and places.* Chicago, IL, US: Center for the Study of Language and Information; New York, NY, US: Cambridge University Press, 1996.

26. Sundar, S., Knobloch-Westerwick, S., and Hastall, M. News cues: Information scent and cognitive heuristics. *Journal of the American Society for Information Science and Technology 58*, 3 (2007), 366–378.

27. Sundar, S., and Nass, C. Conceptualizing sources in online news. *Journal of Communication 51*, 1 (2006), 52–72.

28. Zanardi, V., and Capra, L. Social ranking: uncovering relevant content using tag-based recommender systems. In *Proceedings of the 2008 ACM conference on Recommender systems*, ACM (2008), 51–58.

29. Zhen, Y., Li, W., and Yeung, D. Tagicofi: tag informed collaborative filtering. In *Proceedings of the third ACM conference on Recommender systems*, ACM (2009), 69–76.

30. Zhu, H., Huberman, B., and Luon, Y. To switch or not to switch: understanding social influence in online choices. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, ACM (2012), 2257–2266.