MindMargin: Capturing Reader Engagement with an Article-Adjacent Commenting Platform

Daniel Haehn
Harvard University
Cambridge, MA, USA
haehn@seas.harvard.edu

Sharon Zhou
Harvard University
Cambridge, MA, USA
zhou12@college.harvard.edu

3rd Author Name
Affiliation
Address
e-mail address
Optional phone number

ABSTRACT
Commenting systems are popular throughout many websites. Reading and writing comments can support passive consumption of media by increasing active user engagement in many ways. Information exchange, personal reflection, and lively discussion are only some possible outcomes. We explore how user engagement can be increased by proposing a new commenting system and interface, MindMargin. In contrast to traditional commenting systems where comments are featured below the article, MindMargin presents comments adjacent to its reference media. Users can post and navigate comments and replies on a horizontal infinite scroll. This exposes users to a diverse array of opinions as they read since comments in our system are anchored to relevant areas of the article. Furthermore, we propose a novel metric for classifying and more visibly displaying ”popular” comments. Our metric includes the number of up- and down-votes as well as the post time. This guarantees that new comments are visible even if many popular comments already populate the article. We define three hypotheses to measure increased user engagement that A) investigate if article-adjacent comments generate more comments, B) explore if permanent exposure to comments while reading impacts individual stance towards an issue or opinion in the article, and C) measure the influence of article-adjacent comments on a reader’s impression of the comments. From a user study with 46 young adults, we accepted the second and third hypotheses. We explain the rejection of the first hypothesis as a result of various factors, including the novelty effect. We also discuss how the rejection of our first hypothesis coupled with acceptance of our second and third, in fact, further support our conclusions on increased reader engagement in MindMargin. Based on our initial evaluation and results, we have concluded that the proposed MindMargin system is superior to traditional commenting systems in increasing user engagement while reading articles.

Author Keywords
Comments; anchored; reader engagement;

CHI’14, April 26–May 1, 2014, Toronto, Canada.

DOI string from ACM form confirmation

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces

INTRODUCTION

Motivation
Many websites support the ability to give content-related feedback in the form of comments. Such websites include news pages, media sites, online shops, blogs and social networks. Both reading and adding comments can engage the user in information exchange, personal reflection and lively discussion. In addition, such engagement can impact real world actions like voting, purchasing a product, making choices or participating in a cause.

Traditional commenting systems are featured at the bottom of the content. This vertical structure impedes the reader’s ability to view comments alongside the reference media. These systems also lack a method of referencing sections of the original media. Furthermore, the popular filter mechanism of upvoting popular comments to enhance their visibility risks suppressing newer, and inevitably less popular, comments.

Previous research on annotation and commenting systems has focused primarily on one of three goals: enhancing reader comprehension [6] [8] [4], increasing critical analysis and reflection [7], and encouraging a diversity of viewpoints [2] [5]. Separately, the ends of these goals comprise critical components of reader engagement. The value of this engagement can be seen from how these goals benefit the reader. Enhanced comprehension provides a clearer and more direct stream of knowledge to the reader and adds value to the content of the article. Critical analysis and reflection encourage readers to construct and develop their own opinions from the reading. With access to a broader spectrum of perspectives, readers are urged to consider alternate opinions and as a result, to adapt, strengthen, or clarify their own.

In our study, we measure and take into consideration overall engagement. When engagement as a whole is optimized, readers will have a clearer understanding of the article, develop their own viewpoints as they step back from the article, challenge these viewpoints as they come across other opinions, and feel motivated to take action outside of the scope of the article. Ultimately, greater engagement encourages readers to step away from the article, taking what they have learned with a clearer perspective and engaging more knowledgeably and more openly with the outside world.

Approach

We explore how different commenting models encourage and facilitate engagement by proposing a new system. As opposed to traditional systems, our system **MindMargin** displays comments in a horizontal infinite scroll alongside the reference media. We compare the two systems by quantifying the quality of readers' comments and their follow-up actions. We take into account the ratio of substantial responses to overall responses, instances of referencing the text correctly/incorrectly, and follow-up actions.

We have several hypotheses regarding the comparison of these two interfaces:

1. Readers will generate more substantial comments using MindMargin
2. MindMargin enhances personal reflection due to direct exposure to other people's comments
3. MindMargin increases positive impressions of all comments

Hypothesis 1 is motivated by [7] that found that students engaged more critically with the text when the text was accompanied by annotations. Hypothesis 2 is motivated by [8], in which the comprehension of students reading a text increased with the geographic placement of annotations next to the text. This horizontal structure prevented comments that were out of context and drew students' attention to relevant comments. The authors also found that this increased comprehension encouraged student response and participation. Hypothesis 3 combines the findings of the two previous hypotheses and considers how comments adjacent of the text increases thought beyond the original intended scope of the article.

Contributions

This project includes the following novel contributions:

- a horizontally structured user interface for anchored comments on websites
- a metric to ensure newer comments are still visible beside popular comments while non-constructive comments are less visible
- insights into how comments can enhance user engagement regarding
  1. online discussions
  2. resulting actions (ideally in offline scenarios)

RELATED WORK

Previous research has analyzed how different user interfaces facilitate the gathering of information and the digestion of articles. We compare related work in the field that has motivated our study and approach to the design of our system.

Horizontal Structure and Locality

Having comments appear alongside a text has aided students reading comprehension and encouraged more critical analysis of readings in the classroom. *NB* is a tool that supports collaborative annotation on PDFs [8]. It allows both students and teachers to make comments alongside the text in the appropriate context. This results in less error in the students comprehension and draws the students attention to relevant comments. It also aims to facilitate clearer information exchange in the context of the article. NB improves the clarity and comprehension of lengthier articles in particular.

While the comments on NB appear next to the article, another platform uses anchored comments that highlight the referenced section of the text [7]. This system was also studied with student users who were reported to reflect more critically on the the text with the tool and to be better prepared in argumentation activities. Students also had both qualitatively richer and quantitatively more discussions.

Apart from research on academic instruction, the *Brussell* system uses anchored information in news articles [6]. While not a commenting system, Brussell gathers information across various news sites as a means to provide accurate references and background information to parts of the the article unclear to the reader. By highlighting and interacting with passages in the original article, the reader can request additional background information from different sources. Brussell also suggests questions, such as ”Who is this person?”, for the reader to pose that probe deeper into the understanding of the article.

The geographic location of comments has nontrivial impact on the readers quality of understanding and consequent engagement of the article. As opposed to traditional vertical structures, a horizontal design provides comments that reference the text either directly or in the appropriate context. The optimization of screen real estate through a horizontally planted commenting system motivates our design for MindMargin.

Skewed and Confused: What is Right?

Determining what is important and what is extraneous in a comment is a challenge that all commenting systems must face. Whereas comment placement, as explored previously, affects the overall quality of comments, we must also consider how to increase the visibility of a diverse set of relevant comments.

Opinion Space presents a novel interface as a solution for diversifying comments in a system [2]. Instead of the traditional list format, OpinionSpace employs a visualization technique that reveals the diversity of aggregate comments. Thus, comments are ordered in any particular way and not shown to the reader with the oldest first, the most popular first, or in another skewed manner as is inevitable with a list. MindMargins horizontal structure naturally optimizes screen real estate and prevents clustering of comments by allowing them to spread along the length of the text.

Because opinions in political news articles are frequently biased, another tool, called *Balancer*, accompanies the reader on news articles and provides information on how politically skewed his/her aggregate source of news is [5]. This widget seeks to engage readers in a diverse set of perspectives.

Both OpinionSpace and Balancer reveal that there exists value, as well as a need, in encouraging the consideration and discussion of divergent views online. However, the latter study also points out that people often gravitate to other people and articles with similar views or views that agree with their own. Thus, comments that disagree with the article are likely to be downvoted by the community and become invisible, further discouraging individuals with different views to read the article. In MindMargin, we aim to address this problem by developing a more objective voting system that upvotes a comment based on a diverse set of criteria, rather than a single thumbs-up vote of popularity.

Another study examines the confusion that some even enriching comments present. *Reflect* is a platform that enables users to summarize each others comments in order to enhance comprehension and provide feedback to fellow commenters [4]. This aims to dissipate the confusion and extract the useful information in a comment. With anchored comments alongside the text, MindMargin is designed to minimize confusion by having readers reference the text directly and other readers have context to posted comments.

Invisibility-Disturbance Tradeoff

In optimizing for comment engagement, we must minimize the disturbance that comments cause to the reader as well as maximize their visibility to ensure the readers engagement with them. Thus, there is a tradeoff between how hidden and how disruptive a comment is to the users reading experience.

With traditional vertical platforms, comments are far from the view of the reader as he/she reads the article. They minimize disturbance but also, considering a readers primary task is to read the article, decrease the chance that he/she will read the comments. These platforms also split the engagement of reading an article from the engagement of reading the comments. This gap encourages some readers to skip some or all of the article and to depend on the comments alone for their source of knowledge from the article. In MindMargin, we seek to bridge this division with the horizontal commenting system. Readers are then encouraged to view comments as a companion to the article as opposed to an occupant of a separate niche.

While comments must be visible to draw and enhance reader engagement, highly visible comments can cause disturbance and instead decrease engagement. When applied to designing user interfaces, active pop-up comments have been found to be highly useful to designers, but also somewhat disruptive to their concentration [1]. An interface that has too many moving elements can disturb the reader's focus more than engage him/her in the material.

Encouraging readers to comment also hinges on the visibility of the commenting system as well as its disturbance to the reader. The readers main goal and primary task are generally to read the article passively. The active decision of commenting on the article must be viewed as secondary. One study showed that readers are more likely to engage in a secondary task if the threshold of that secondary engagement is low [3]. However, the methods that lower barriers to the task most effectively, like pop-ups, are also highly disruptive. In the design of MindMargin, the horizontal structure inherently adds to the visibility and disturbance of the system when compared against a vertical platform. Thus, to optimize for engagement, we must focus on actively minimizing the systems disruption.

APPROACH

We implemented two commenting systems as classical client/server web-applications. The first commenting system is MindMargin with anchored comments on a horizontal infinite scroll next to the reference medium. The second commenting system is a traditional vertical interface. Users interact with the clients as a front-end system using a web browser. The clients communicate with the server back-end using AJAX to a) request existing data or b) persist new data. The server reads and stores data in a relational database.

Front-end

The client interfaces consist of clean user interfaces to avoid design clutter and distraction. Figure 1 shows the MindMargin system in action. The application is split into two sides: The reference media on the left and and an adjacent commenting system on the right. The commenting system displays comments in a horizontal infinite scroll. Thus, an unrestricted amount of comments can be linked to the reference media. Navigation within the infinite scroll component can be performed via mousewheel interaction (either left/right or top/down scrolling with the same effect) or by adjustment of a slider on the bottom of the right split screen. Comments are anchored to the horizontal reference point of the media by thin dotted lines. If a comment has replies, a dropdown button appears on the comment's footer. Lighter in color, replies to comments appear vertically under their comment when the button is clicked. This arrangement optimizes horizontal real estate by reserving horizontal space for parent comments. Finally, while navigating through the infinite scroll, the reference medium remains fixed on the left for quick reference against referential comments and replies.

We have also implemented a metric to distinguish between popular and regular comments that appear separately. For greater visibility, popular comments are displayed directly adjacent to the article in the MindMargin interface and directly under the text, or first, in the traditional layout. Upvotes and downvotes indicate and impact comment popularity. We also consider comment time to determine individual comment relevancy and the order in which they appear. This is because new comments are at a temporal disadvantage in garnering votes compared to their older counterparts. Comments posted within the last 60 minutes are labelled as popular, regardless of their vote number. This prevents new comments from losing visibility, especially when a large number of comments exists. We define our metric for popularity as follows:

$$c_{popular} = \begin{cases} v_{up} + T & \text{, if } v_{up} > v_{down} \\ v_{up} \cdot 1.3 - v_{down} + T & \text{, if otherwise} \end{cases}$$

Figure 1. The MindMargin system consists of a web client (shown here) and a server side back-end.

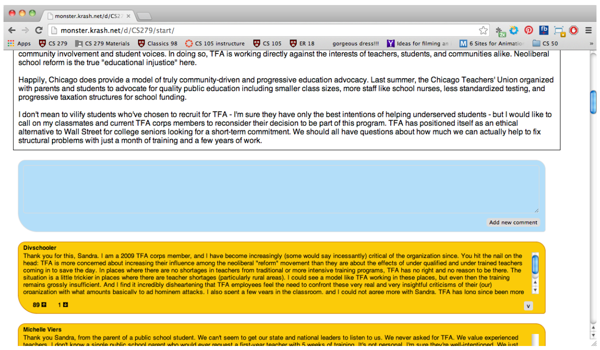

Figure 2. The traditional comment system with a vertically ordered design.

$$T = \begin{cases} 9 & \text{, if } c_{time} < 60\text{minutes} \\ 0 & \text{, if otherwise} \end{cases}$$

Comments are marked popular if $c_{popular} > 8$. This metric was included in both commenting systems.

Figure 2 shows the traditional vertical commenting system. The reference media appears first and on top of the commenting system that follows below. Navigation within the article as well as within the comments can be performed via top/down scrolling. The organization of replies and up- and down-voting is similar to the MindMargin prototype.

The front-ends of both prototypes were written in JavaScript using the popular jQuery and jQuery UI libraries. A model-view-controller pattern was chosen to structure the application code base. The user interface itself were created to adjust responsively to any window size.

Back-end

The server component of MindMargin and the traditonal system was written in PHP and communicates with the front-end and a relational database. Communication with the front-end is ensured by providing a REST-API which can be called via AJAX. The entity model of the client is replicated on the server. Data is read and stored using a custom and fully generalized object-relational-mapper. We chose MySQL for our database with a table each for users and comments.

Throughout the implementation, we followed an iterative approach to programming our software. The developer team was small. All developed code is released under the BSD open source license on github.

Figure 3. The original article from The Harvard Crimson regarding Teach for America (TFA) which was chosen as reference media in our user study.

EXPERIMENT

Here we describe our experiment. We performed a blind user study on young adults. Participants were randomly associated to either MindMargin or the traditional commenting interface.

Participants

We initially recruited 106 online participants for our user study and evaluation, of which 46 completed the study (30 female). Participants were college students, aged 18 to 25. The reading frequency of news articles among participants ranged from daily to almost never.

Experimental Conditions

The two conditions in our study were MindMargin and the traditional vertical interface, seeded with 39 existing comments from a news article. We selected a recently published article from the Opinion Section of The Harvard Crimson, titled Dont Teach for America. as seen in figure 3. We chose this article on the basis of its opinionated nature and its relevance both in recent news and to our anticipated participant pool. The article already had over fifty comments, from which we selected the top 39 as ranked by Disqus, the existing commenting system in The Harvard Crimson, to be used in our study. The same comments were used in both conditions. In the traditional vertical interface, they appeared in the identical order as ranked in the original article. In MindMargin, they were anchored to the article based on textual references in the comment.

Design and Setup

We employed a between-subjects test, with participants assigned randomly to one of the two conditions (19 MindMargin). In order to reproduce the conditions under which one would normally read a news article, we chose to recruit participants online and self-select themselves into reading the article of interest. In order to motivate our participants to actually read or skim the article, instead of skip it, we chose not to use monetary or other time-sensitive incentives. Instead, we chose to design the experiment around the survey question, ”Do you (really) think like a Harvard student?”. Figure 4 shows the landing page of our experiment. We then asked participants follow-up questions to verify that they read the

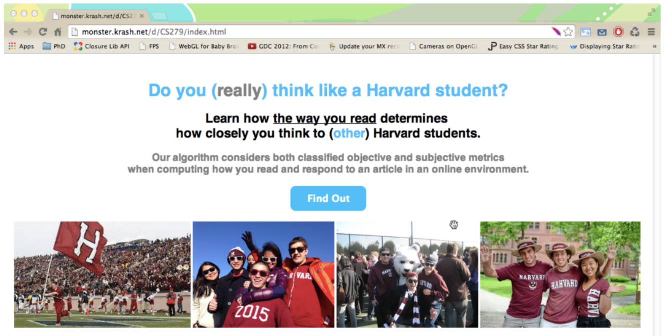

Figure 4. The landing page of the performed user study.

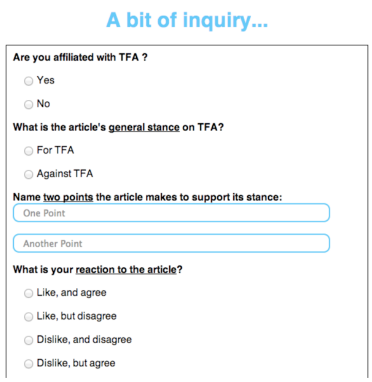

Figure 5. Parts of the post-experiment questionnaire.

article, which included both the overall stance of the article and two pieces of supporting evidence used in the article. Figure 5 shows the post-experiment questionnaire. Participants were not allowed to refer back to the article once provided the questionnaire.

Procedure

Participants were recruited online through social media and college listservs. They were given an initial questionnaire asking basic demographics and reading frequency. Before given the article, they were also asked to provide a username or pseudonym, or to remain anonymous. During the reading of the article, participants were allotted 10 minutes. After 2 minutes, they were permitted to proceed. The 2-minute delay was to ensure reading of the article, but did not seem to prevent fast readers from proceeding as the average reading time was 3 minutes 47 seconds. In the follow-up questionnaire, reading verification questions were first asked of the article. Participants were also asked their personal stance on the article, whether they liked the article, and whether they agreed with the article. They were also asked to self-report whether they read the comments in the article and to provide two adjectives that described either their reaction to the comments or a description of the comments.

RESULTS AND DISCUSSION

In this section, we report on the findings of our user study. Overall, we observed an increase in user engagement when comparing the proposed MindMargin interface against the traditional vertical commenting system. We were able to accept two of the three initially defined hypotheses.

Our first hypothesis predicted an overall increase in newly generated user comments, replies, upvotes, and downvotes when using MindMargin. We believed that having the comment system permanently visible on the side while reading an article, as in MindMargin, should increase these actions. However, this was not the case: we **rejected Hypothesis 1** since we did not observe an increase in the reading rate of comments or in interactions with comments when using our system. In fact, the participant group using MindMargin stated a 73% reading rate of comments while the control group using the traditional interface stated a 85% reading rate. In addition, there was no statistically significant increase in commenting rate, reply rate or the rate of up- and downvoting. We have several theories to explain these phenomena. First, we think that the users were overwhelmed by the number of existing comments. We populated both interfaces with 39 comments and replies, extracted from the original article. These comments covered a wide range of controversy regarding the content and participants might have thought that everything was already said. Second, since we were testing a rather unusual interface, we think that the novelty effect of introducing a new horizontal interface should not be underestimated. Third, because the study was hosted on an external site from that of the original article, which typically does not display news, participants may have concentrated more on reading the article than providing comments. Fourth, we do not know the number of comments that participants who reported to have read the comments did actually read.

Our second hypothesis was defined as an increase in personal reflection when using MindMargin. Exposure to a range of controversial comments should result in the rethinking and revising of ones own opinions. All participants were asked their stance, from Strongly For TFA to Strongly Against TFA, on a Likert scale. Using data from participants who reported to have read the comments (see above), we observed the percentage of participants who claimed a strong stance on the article. Of those assigned to the MindMargin interface, only 16% reported to be either Strongly For TFA or Strongly Against TFA. In contrast, 26% of the participants using the traditional commenting system reported either extreme stance. The distribution of the Likert values is also normal for MindMargin and a U-shaped curve for the traditional commenting system. We performed a Shapiro-Wilk normality test ($alpha = 0.05$) on both distributions. MindMargin rejects the null-hypothesis with $p = 0.1306$ and therefore is normally distributed. The traditional prototype accepts the null-hypothesis with $p = 0.0205$ and is therefore not normally distributed. We created Normal Q-Q plots for both (figures 6 and 7). This reveals that despite no increase in the rate of reading comments, the MindMargin interface was able to encourage users to consider other opinions and viewpoints. This suggests a greater user engagement with the comments with the MindMargin interface. Therefore, we have **accepted Hypothesis 2**.

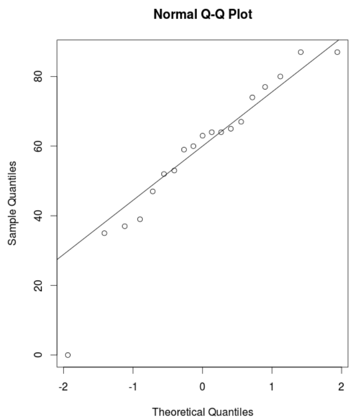

Figure 6. The personal stance distribution for MindMargin participants (normally distributed according to the Shapiro-Wilk normality test with $alpha = 0.05, p = 0.1306$).

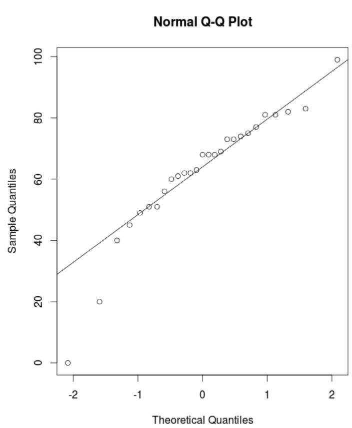

Figure 7. The personal stance distribution for traditional interface participants (normally distributed according to the Shapiro-Wilk normality test with $alpha = 0.05, p = 0.0205$).

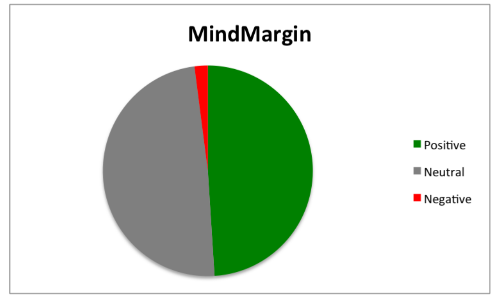

Figure 8. When using MindMargin, the majority of participants described the comments as positive.

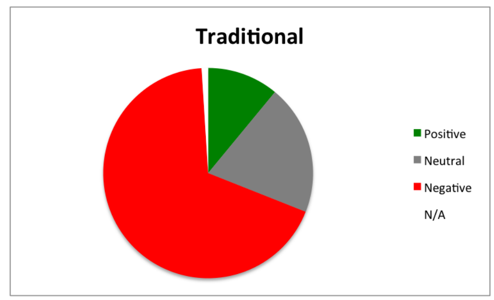

Figure 9. The majority of participants using the traditional commenting system described the comments as negative.

Our third hypothesis predicted an overall increase in positive impressions on comments when using the MindMargin interface. We asked participants who read the comments to input two adjectives in free-text describing either their reaction to the comments or a description of the comments. We then classified these adjectives using a three-bin classifier (Positive, Negative, and Neutral). Positive was assigned to positive reactions to comments, such as interesting, well thought-out, and engaging. Negative was assigned to negative reactions to comments, such as annoying, useless, distracting. Neutral was assigned to descriptive input about the comments, such as long and subjective. Finally, a few outliers, such as trolls and whatever, were removed. We observed a drastic change of impressions when using MindMargin. As seen in figure 9, the majority of participants using the traditional commenting system described the comments as negative (68%). In contrast, when using MindMargin, the majority of participants described the comments as positive (48%) or neutral (48%) as seen in figure 8. Outliers were also observed only to occur in the traditional commenting system. We have therefore **accepted Hypothesis 3**.

The increased positive reaction to comments suggests that users perceive comments to be more substantial and worth greater consideration in the MindMargin interface. Even though participants were exposed to the identical comments and the comments were initially written for the traditional vertical interface, reactions to the comments were significantly more positive when using MindMargin. The rejection of Hypothesis 1 renders this finding even more surprising because there was no statistical difference in the number of people who read comments in either interface. The rejection ultimately supports the increase in user engagement with the

MindMargin interface, and even suggests that MindMargin is not a distracting interface since participants can still opt out of reading comments if they so choose. Moreover, those who did not opt out of reading the comments found the exposure to anchored, relevant comments in MindMargin to be considerably more engaging and less distracting. In addition our quantitative results, we would like to quote qualitative feedback from a MindMargin user, suggesting actions he/she took beyond the scope of reading and commenting article: This article showed me a new perspective on TFA, which after doing research, I have realized I agree with. No feedback suggesting actions outside the scope of the article was received from participants with the traditional commenting system.

CONCLUSION

In this paper, we studied how commenting systems can increase user engagement with comments while reading articles. We report a correlation between the novel MindMargin interface and increased personal reflection as well as a more positive overall impression of existing comments on the article. The key difference between traditional commenting systems and MindMargin is that in the latter, comments are anchored to specific passages of the reference media and are placed on a horizontal infinite scroll. We developed two commenting systems, one using the traditional vertical interface and the other using the MindMargin interface. Then, we performed a user study for evaluation. Our key findings include that being exposed to relevant comments during reading increases personal reflection. This results in 10% less extreme positions regarding the context of the reference article. Additionally, the overall impression of comments significantly diverges. Users of a traditional commenting system report 68% of comments to be negative, while users of MindMargin report only 2% of comments to be negative. Contrary to our initial hypothesis, MindMargin does not generate more comments or more replies. We think that this is primarily due to a) the novelty effect and b) the existence of previously seeded comments. Therefore, future research will include a user study without already seeded comments. In addition, we plan to expand the participant pool to include participants of all ages. Finally, we plan to pursue research on a commenting system like MindMargin, but for videos and music, that anchors comments to certain times or time-intervals within a given recording. Research into annotations on visual pieces, other than text, is also being considered.

REFERENCES

1. Ericsson, M., Baurén, M., Löwgren, J., and Waern, Y. A study of commenting agents as design support. In *CHI 98 Cconference Summary on Human Factors in Computing Systems*, CHI '98, ACM (New York, NY, USA, 1998), 225–226.

2. Faridani, S., Bitton, E., Ryokai, K., and Goldberg, K. Opinion space: a scalable tool for browsing online comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, ACM (New York, NY, USA, 2010), 1175–1184.

3. Hoffmann, R., Amershi, S., Patel, K., Wu, F., Fogarty, J., and Weld, D. S. Amplifying community content creation with mixed initiative information extraction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, ACM (New York, NY, USA, 2009), 1849–1858.

4. Kriplean, T., Toomim, M., Morgan, J., Borning, A., and Ko, A. Is this what you meant?: promoting listening on the web with reflect. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, ACM (New York, NY, USA, 2012), 1559–1568.

5. Munson, S. A., Lee, S. Y., and Resnick, P. Encouraging reading of diverse political viewpoints with a browser widget. In *ICWSM* (2013).

6. Wagner, E. J., Liu, J., Birnbaum, L., and Forbus, K. D. Rich interfaces for reading news on the web. In *Proceedings of the 14th international conference on Intelligent user interfaces*, IUI '09, ACM (New York, NY, USA, 2009), 27–36.

7. Wolfe, J. Annotations and the collaborative digital library: Effects of an aligned annotation interface on student argumentation and reading strategies. *I. J. Computer-Supported Collaborative Learning 3*, 2 (2008), 141–164.

8. Zyto, S., Karger, D., Ackerman, M., and Mahajan, S. Successful classroom deployment of a social document annotation system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, ACM (New York, NY, USA, 2012), 1883–1892.