Deployment of a Collaborative Multimodal Annotation System for Instructor Feedback and Peer Discussion

Dongwook Yoon[1], Nicholas Chen[2], Bernie Randles[1], Amy Cheatle[1], Corinna E. Loeckenhoff[1], Abigail Sellen[2], and François Guimbretière[1]
[1]Cornell University and [2]Microsoft Research

ABSTRACT
In educational environments, digital annotation tools have long been used to facilitate collaboration over documents, but their communication capabilities often lack the expressiveness and richness of face-to-face conversation. Recently, RichReview, a collaborative, multimodal annotation system, was proposed as a way to bring more of the benefits of being face-to-face contact into remote, asynchronous interactions. We conducted a series of field deployments to evaluate how classroom activities could benefit from using RichReview. We first deployed the system to support instructor feedback in an undergraduate course. In a second study, the system was used for online peer discussion about weekly readings in a graduate level course. Throughout a series of iterative evaluation cycles, we were able to improve the system by introducing several new features that facilitated collaborative commenting. The results of the studies revealed (1) that the waveform indexing feature facilitates consumption of the instructor's voice comments, (2) that the annotation anchoring feature encourages the generation of more specific and detail-oriented comments, and (3) what motivates or demotivates the use of voice comments in a classroom setting.

Author Keywords
Multimodal input; collaborative annotation; anchored comment; instructor feedback; online peer-discussion.

ACM Classification Keywords
H.5.3 Group and Organization Interfaces: Collaborative computing; H.5.2 User Interfaces: Interaction styles; H.5.1 Multimedia Information Systems: Audio input/output.

INTRODUCTION
As remote collaboration tools become an increasingly prevalent feature of both online and offline classrooms [4,9,17,39], a key challenge facing these tools is to bring the richness and effectiveness of face-to-face communication into digital practices. Researchers in the fields of CSCW and CSCL have explored digital tools to supplement in-person communication [12,31]. Among them, Yoon et al. presented the RichReview annotation system [38] to effectively facilitate remote conversation over a shared digital document. To this end, RichReview combines ink, voice, and gestural annotations anchored and presented alongside the text of the document. It also offers a visual representation of voice annotations, to facilitate the review and re-review of audio comments.

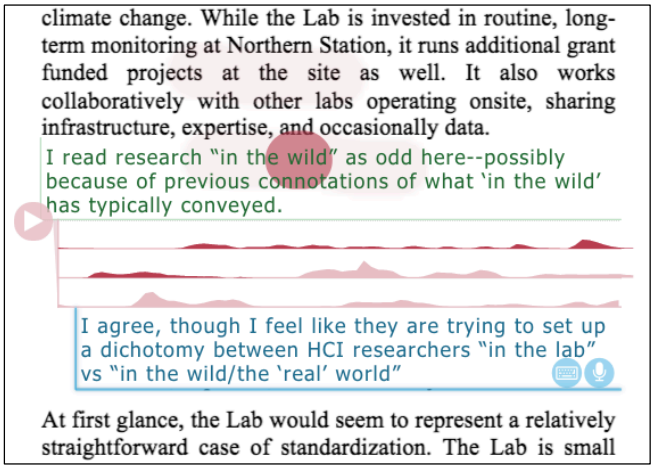

Figure 1. RichReview[++] screen shot. Multiple users created a threaded conversation using a mixture of text, voice, and pointing gestures.

Although a small-scale formative laboratory study validated the interface concept [27,38], the utility of RichReview has thus far been unproven in real world settings. The present paper addresses this gap and describes the deployment of a newer, cloud-based, version of RichReview (i.e., RichReview[++]) in two different educational contexts: 1) to support instructors delivering term paper feedback to a small undergraduate seminar; and 2) to support reading assignment discussions for a graduate student class.

This series of evaluations was designed to answer the following questions:

- What technical specifications and adjustments are required to make the RichReview^{++} system practical and usable in the educational settings?
- Which interface features of RichReview are perceived to be most useful by students, and which aspects need further improvement?
- What can we conclude about the potential for multi-modal tools to add more richness and expressivity to asynchronous, remote communication? What is their potential role in future educational settings?

Our first deployment examined instructor feedback on a writing assignment in an advanced undergraduate class. Although this pattern of use was similar to the design goals of RichReview [38], the original tablet-based system did not reflect the realities of general user populations in classroom where the majority of students use laptops. To address this problem, we implemented a desktop web viewer to display annotated RichReview documents via a standard web browser to student users. We found that students rated multimodal feedback favorably and sometimes even considered it as a proper substitute for meeting a professor in person during office hours. Analysis of the logs and student feedback also suggested that the visual waveform indexing of voice annotations successfully supported initial playback as well as replay of audio comments.

Our second deployment focused on peer discussion, another instance of document-centric collaboration in classrooms. Peer discussion is known to be beneficial for strengthening students' mental models of the topic [32] and broadening their perspective on problem solving [22]. However, little was known about the potential benefits of adding multimodal features to online discussion. In contrast to the first evaluation, where only the instructor created annotations, this setting required every student user to contribute annotations. To achieve this, we implemented RichReview^{++}, a significantly enhanced version of the system that allowed students to use their own laptops to annotate a PDF document using text, audio and highlighting. Results suggest that RichReview^{++} has pedagogic affordances that encourages socially driven learning processes, but also highlight some requirements for a wider adoption of the system. These include better support for the revision of voice comments, and appropriate instructional designs such as assignment timeline alignment.

This paper additionally describes the extensive changes we made to the RichReview system to support publically shared comments, annotations spanning different documents, text annotations as well as enabling the tool to run on virtually any computing device.

Our research presents new insights into how multimodal comments anchored to the text and replayed as a continuous stream can facilitate online communication in real educational settings. Thus, it provides critical insights into the specific user characteristics and contextual features that need to be considered when designing multi-modal annotation platforms for different educational settings.

RELATED WORK

Online Tools for Classrooms

Many collaborative tools have been developed in support of educational settings. These tools include discussion forums [13], dedicated feedback interfaces [2,16,38], and various forms of text-based commenting systems [4,9,39]. Although classroom activities are extremely diverse, our current work focuses on the use of multimodal document annotations to support two ubiquitous classroom activities: instructor feedback to students and peer discussion among students.

Enhancing Instructor Feedback

Receiving feedback from an instructor is an essential pedagogical interaction [30]. As it stands, instructors have many choices on how to deliver this feedback. Examples include word processor comments [37], ink [3,16], audio files [5,23], anchored audio comments [24], and even video recordings [8,31]. Two major trends have emerged from the literature. First, the use of non-textual modalities, such as voice and inking, has been found to be more engaging [24], and provides a new dimension of interaction between instructor and students [3,16]. Secondly, anchoring comments over the students' submissions makes the feedback easier to understand [37] and more engaging [8].

Supporting Peer Discussion

Online forums and message boards are a popular means of facilitating discussion in the classroom. A subset of this type of software are tools that enable anchored discussion: situating conversation between peers in context of the document that is being discussed [4,7,9,10,13,39]. Previous research has found that anchored discussion increases collaborative learning by focusing students' attention on specific content [1], and promoting constructive discussion [26]. Our deployment extends this work, by exploring how the anchored discussions featuring alternative modalities can facilitate social interaction and the learning process.

Multimodal Annotations

Different annotation modalities have unique strengths. Typewritten text, which is widely used, is fast to write and read. Freeform inking [14,29] allows for lightweight markup and the use of abstract symbols. Audio can better communicate nuanced ideas and emotions [21,25]. Pointing and other gestures can provide a point of reference in the mediated conversation [34,38]. Combining inking with voice input [19,33,35,36] not only visually augments the spoken comment, but also offers ink marks as navigational cues for indexing voice recording.

RichReview [38], upon which the system we deployed is based, presented a method to integrate different annotation modalities into a coherent stream within a single document.

Laboratory studies found that RichReview was effective as a document-centric communication tool in that it was capable of embedding collaborative conversations within the document itself [27,38]. This paper extends these past findings to real-world educational environments.

THE RICHREVIEW SYSTEM

The starting point for the system we deployed was RichReview [38], an annotation system that combines ink, audio and gestural annotation to provide a face-to-face like communication experience for collaborators reviewing and providing feedback on documents. In addition to voice and inking, the tablet-based system employed pen-hovering as a deictic gesture for referencing as a support for spoken comments. Also, RichReview facilitated accessing, skimming, and browsing of multimodal comments by providing quick indexing to the recorded data stream (e.g., clicking over a waveform visualization to jump around in an audio recording).

Although many of the features of RichReview were applicable to classroom activities, the fact that RichReview was not specifically designed for educational settings necessitated a number of enhancements (e.g., web-based access, the option to create private annotations) that were refined through an iterative design process. We now discuss each of the two deployments in more detail.

DEPLOYMENT #1: RICH INSTRUCTOR FEEBACK

The goal of the first deployment was to evaluate the utility of multi-modal annotations in the context of instructor feedback on a term paper. A particular focus in this regard were students' perceptions of the effectiveness of the feedback but we also considered their emotional responses and the interpersonal dynamics entailed in this approach.

A Web-Based Viewer for RichReview

In the original system, making full use of the capabilities of RichReview required specialized hardware (Windows tablets). While it would have been possible to provide each student in the small class with such a device, it was inherently impractical and non-scalable solution considering deployment costs. Moreover, given that most students were highly reliant on laptops, requiring students to carry around an additional device for a single task seemed to be overly burdensome.

Our solution was to create a new, web-based viewer for documents annotated using RichReview. Thus, the instructor can use a tablet computer to create her comments (leveraging the full feature set of the system) and then upload the annotated document where students can access using the browser on their device of choice via a URL the instructor provides. Figure 1 shows a screen capture from the web viewer. Our implementation was hosted on a cloud service (Microsoft Azure) and satisfied the security and accessibility requirements of the Family Educational Rights and Privacy Act (FERPA) in the United States for educational electronic data storage and handling.

Deployment Procedure

The system described above was deployed in an undergraduate level psychology course in the Fall 2014 semester at our university. Students in the class submitted term paper drafts as PDF files via email. The instructor commented on each paper using with digital ink, audio, and gestures before uploading the results to a cloud-based web service. Students received feedback 10 days before the final version of the paper was due.

Participants

Participation was voluntary and students could choose between receiving handwritten notes on a printout and RichReview-based comments. Exit surveys were also optional and the instructor did not have access to student responses. The class was composed of one instructor and 17 students, 16 of which participated in the study and 13 of which answered the exit survey. The participating students were nearly all female undergraduates (mean age: 21.0); one student was male, and one was a graduate student. All students used laptops to view the feedback provided.

Analysis

We monitored and logged students' online activities such as page navigation and voice playback. An online exit survey was conducted after the final paper submission. The survey collected subjective ratings in a 5-point Likert Scale, and gathered qualitative responses as textual descriptions with a keen focus on usability of the various system features, and preferences to different ways for giving instructor's feedback (e.g. "I prefer RichReview to having an office hour meeting"). The students received paper-based handwritten feedback from the same instructor earlier during the semester creating a baseline for students to compare.

Results

Essays ranged from 10 to 22 pages in length (M = 15.7 pages, SD = 2.73). The instructor made digital ink markups for typographical edits as well as voice comments for detailed commentaries. On average, the instructor made 51.4 voice comments (SD = 11.7) per student essay with a mean length being 14.9 sec (SD = 15.1). Students found the system easy to learn (Learnability: M = 4.60, SD = 0.52) and used the audio recordings extensively: 95.4% of voice comments were replayed by students, and many of them were replayed multiple times (see below for details). The general user response was very positive: one commented that "This was by far the best experience I've had while revising a paper (P10)". Students reported a willingness toward continued use and recommended its use to peers. To quote P6, "Would definitely use again and would recommend this to others!"

Benefits of Waveform Indexing for Audio Consumption

One recurring theme of the qualitative feedback was that the easy-to-use audio replay was helpful for consuming the recorded comments. We analyzed logs of online activities to take a deeper look into replay patterns. The results showed that audio re-listening was very popular. As shown in the replay count histogram of Figure 2, the majority (73.1%) of

the voice comments were replayed more than twice, a few of the comments (6.44%) were replayed more than 10 times, very extensively.

Deeper analysis of data revealed that waveform indexing was the key facilitator of the re-listening behavior. More than half (8) of the students used the waveform indexing feature more than 10 times. Across all playbacks, the play button was clicked in 81.7%, waveform indexing was used in 12.4%, and gesture traces were clicked in 4.6%. We also found that the number of replay counts via indexing waveform for a given voice comment was significantly correlated with the comment's length ($N=823$, $r=0.22$, $p<.001$). However, a similar pattern did not emerge for the play button. We believe that this suggests that waveform indexing was preferable for dealing with long audio comments. This finding was re-enforced by our survey. Participants ratings suggest that the feature was very helpful for understanding audio comments ($M = 4.63$, $SD = 0.52$, 8 responses). Participants reported that they used this feature to skip to a specific section (P6) or repeat a missed word (P4).

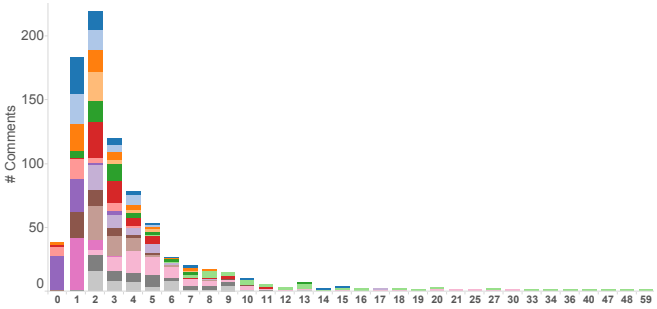

Figure 2. Distribution of replay counts. Colors indicate different users. Most comments were replayed more than twice. Also note that a few comments were totally missed.

Multimodal Annotations vs. Longhand Comments

Participants rated RichReview annotations more highly than written comments for receiving feedback about writing issues related to factual content ($M = 4.27$, $SD = 0.79$) and structure ($M = 4.55$, $SD = 0.52$). There was no preference between the two methods for comments pertaining to grammatical errors and typos ($M = 3.09$, $SD = 1.45$). This result is consistent with Chalfonte et al.'s finding that voice is superior to text when describing higher-level (structural, or semantic) writing issues in comparison with local problems [5], albeit in an educational context.

Multimodal Annotations vs. Office hours

Surprisingly, the majority (11) of participants preferred RichReview annotated documents over office-hour meetings ($M = 3.91$, $SD = 1.22$), and believed that they offered an acceptable substitute for in-person meetings ($M = 3.91$, $SD = 1.04$). Qualitative comments offer two explanations for this result. First, students wanted to "incorporate all the comments" and make sure that they were "doing everything that the instructor suggested" (comments from P4). RichReview was useful because recorded comments were hard to miss and could be tackled one-by-one. Also, the rapid stream of feedback received when meeting in person made students worry about whether they were missing or misunderstanding the instructor's comments. In contrast, the recorded comments could be replayed multiple times when they were not clear. To quote P4, "I can listen to everything multiple times if I didn't get it which also made it less intimidating". On the other hand, a few students (2 of 13) thought that in-person meetings offered a more immediate interactive dialogue which RichReview's asynchronous interaction did not offer.

To summarize, students confirmed that RichReview was a useful channel for delivering feedback from an instructor. The student's ability to efficiently consume the instructor's voice comments appears to be a primary factor. Fast and easy access to audio content played a particularly important role in this task.

DEPLOYMENT #2: SUPPORTING PEER DISCUSSION

The goal of the second study was to determine whether the situated multimodal annotations RichReview influences discussions about reading assignments among students.

RichReview^{++}

For this second deployment, participants were required to both produce and consume comments. This led us to implement a new, fully web-based version of RichReview which we call RichReview^{++}. As in the first deployment, our goal was to make the system accessible to anyone with a basic laptop (equipped with a microphone) running a modern web browser such as Google Chrome. Besides the engineering effort required to capture and process audio from inside a web browser, the main focus of our effort was to re-frame the interface from a pen-centric approach to a keyboard/touchpad centric approach.

The first major change to classic RichReview was the addition of text as one of RichReview's core annotation modalities. This means that typed comments behaved in a similar manner to ink and audio annotations in the system. In order to retain a consistent and straightforward interaction for creating a comment regardless of the user's modality choice we designed a set of keyboard-centric interactions. In this set of interactions, the 'Enter' key was a universal shortcut for creating a new comment. To create audio comments, users press 'Enter' on its own to begin recording ('Enter' + Speak). But if users continue typing after the 'Enter' key is pressed ('Enter' + Type), we terminate the voice recording and create a new text comment instead.

Private Annotations

Although peer discussions took place around public annotations, the literature suggested that support for private annotations would important as a staging step before creating public comments and also to support active reading [20]. Consequently our system added the capability for users to create personal annotations in the system. Our system supported two types of private annotations: textual notes and highlights. We thought it unlikely that the users would

record voice comments for themselves, therefore we did not implement private voice comments. There were three principles that went into the design of the private annotation feature. First, public and private annotations should be easily visually distinguishable. Second, the creation of private annotations should require minimal effort. Lastly, a private annotation should be easily moved from a private to public status (and vice versa).

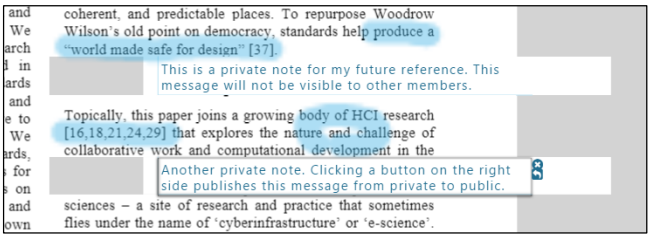

Figure 3. Private Notes and Highlights. Private notes are extended across the page boundary.

In the tool, users create private annotations using the 'Ctrl' modifier key in conjunction with the regular annotation shortcut. For example 'Ctrl' + 'Enter' creates a new private text comment and 'Ctrl' + Cursor-Drag creates a private highlight. A private annotation is also placed in-between text lines like public comment, but extends over the page bounds to create a clear visual contrast (see Figure 3). Converting a private comment to a public one is accomplished by clicking a 'publish' button on the corner of each text box.

Comment History for Browsing and Re-visitation

To enable quicker and more complete browsing of comments we added a Comment History. The Comment History displays a series of chronologically sorted links that lists the full set of comments present in a document. Each comment is represented as a small icon and when clicked changes the view to show that particular comment and centers and centers it in the viewport.

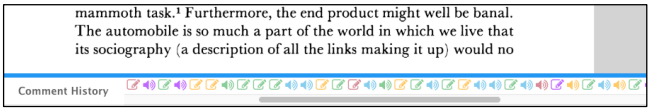

Figure 4. Comment history feature. A user can click one of the chronologically sorted links to existing comments to jump to the relevant page and the selected comment is highlights.

Deployment Setting

The RichReview^{++} online discussion system was deployed to a graduate level social science course in the Spring 2015 semester. In the class, an instructor taught 18 students, with 4 of them connecting from a satellite campus to the main class via a videoconferencing system. The class centered on individual readings and class-wide peer discussion activities. Assigned readings for a given week were first discussed online and then in a class session lasting 2.5 hours. Online discussion contributions made up 20% of the grade. The reading materials were composed of either 4-5 different conference papers, 3-4 chapters of a textbook, or a mix of both, totaling 150-250 pages per week. The classroom discussion was open-ended and student-directed; with the lion's share contributed by three discussion leaders designated by weekly rotations.

Procedure

We compared our system to the BlackBoard Learning Management System (hereinafter referred to as BlackBoard), a commercial system which is extensively used at the university. The students first used BlackBoard for one week, then used RichReview^{++} for two weeks before reverting to Blackboard for one week (Table 1). In response to this initial evaluation, we made improvements to the system, and conducted a follow-up study that focuses on evaluating the new features.

Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12
BB	RR	RR	BB	System Improvement		RR	RR
Main Study						Follow-Up	

Table 1. The semester schedule of the peer discussion study. BB stands for a BlackBoard Week, and RR for RichReview^{++}.

Measures

At the end of each sub-block of the study, we conducted an interview and survey with questions about the discussion tools. The 30 minute interviews were semi-structured and open-ended. To draw out emerging themes, the lead investigator analyzed the interview data via 2-step encoding—open coding followed by flat coding. The format of the surveys was similar to that of the previous instructor feedback study (Likert scale ratings + open-text comments). We also logged and analyzed online user activities, including audio replays and commenting.

Participants and Grouping. Participant information and informed consent procedures were similar to the first evaluation. Fourteen of the 18 students participated in the study. The participants were mostly graduate students (1 undergraduate) in their mid-20s (M = 26.5, SD = 3.5). Four of them were female. Their knowledge levels regarding the course topic spanned a wide range, from novice to very familiar. Five Participants were native English speakers, and the others were ESL students whose English proficiency ranged from intermediate to fluent. We placed the 14 participants into 3 groups with 4-5 members each, because having all participants commenting in a single group could easily clutter the document space with RichReview^{++}'s visually rich comments. Groups were balanced regarding background, gender, English proficiency, and campus location.

Result from the Main Study

Participants successfully used the two discussion systems for their desired online activities. Participants reported that the user interfaces of both systems were intuitive and fluid enough to support the online discussion task without any significant difficulties. Even though it was the first time the participants used RichReview^{++}, they were able to navigate its user interface with ease. Some participants reported minor

difficulties with the page-based navigation for quick skimming, but that did not hinder the constructive use of other system features. After the first block, 11 participants were interviewed and 13 participants replied to the online survey. Below, we present emerging themes extracted from the interviews, and support them with survey and log data.

Annotations versus Comments

At a high level, the most striking pattern that emerged was that users treated Blackboard as a commenting tool and RichReview^{++} as an annotation tool. We saw this pattern both in quantitative and qualitative measures. On the quantitative side, we examined the size and number of entries generated in each system. BlackBoard posts were long (Mean word count = 1159.4, SD = 453.5) essay style notes, which were often followed by short replies (Mean word count = 69.5, SD = 23.6). There were 19 post threads and a total of 4 short replies throughout a two week time period. For RichReview^{++} we observed 251 textual comments and 90 voice comments. A total of 322 threads were formed with most of them (306) being singular without a reply (Mean word counts = 31.8, Mean recording duration = 19.6). Sixteen comments had replies, but thread length was short (2 or 3) without much back and forth conversation. Three participants created the majority of their comments using the recording feature, 3 others made mixed use of modalities, and the rest created their comments mostly using typewritten entries.

On the qualitative side, students reported that working with Blackboard prompted them to produce comments that involved more synthesis i.e. combining concepts across different pieces of text (P4, 8, 11). In contrast, the embedded approach offered by RichReview^{++} led to more targeted comments, focusing on low level ideas. This led to a style that was more conversational (P1, 8, 11) and easier to understand (P4, 12, 14). In particular, participants noted that it was much clearer what a given comment was referring too (P7, 12, 14).

One advantage of anchored comments that emerged was that students could deepen their understanding of the texts via socially driven learning [10]. As they read the passage in RichReview^{++}, students paid more attention to the part that had a comment located underneath it (P7, 9, 11, 14) [20]. Some participants liked to quickly browse through all the comments before starting a reading in order to get a sense of which part of text was deemed interesting by other students (P5, 9, 12, 14). In particular the Comment History was rated to be very useful (M = 4.36, SD = 0.5). The feature was used as a summary of peer activities. For example, students used Comment History to quickly browse through existing comments to get a sense of how other students interpreted the readings (P4, 9, 11, 13) or to see how others' ideas have developed in time as they were reading the texts (P9, 12). Additionally, P3 could spot the 'hottest' article by counting the number of icons in the comment history.

Benefits and barriers of voice

Participants told us that they liked listening to voice comments (M = 4.08, SD = 0.67). Voice comments provided a more personal, nuanced feedback and made it easy to identify the author of the comment. However, participants identified several obstacles blocking more widespread adoption of audio annotations. Participants (5 of 11 interviewees) told us that they sometime had difficulty skimming through voice recordings, but most of the concerns were raised about the production side. Participants found that it was harder and more time-consuming to create voice comments (M = 3.08, SD = 1.38) than to create text comments (M = 4.33, SD = 0.49). Four major reasons stood out. First, because our system did not implement the speech editing feature available on the tablet software, entire recordings needed to be re-recorded when there was a mistake. Second, since people didn't want to introduce long pauses in the middle of voice comments, they felt compelled to keep speaking, which interfered with their thinking (5 of 11 interviewees). As a result, students felt that they were less successful communicating their ideas via voice comments (M = 3.17, SD = 1.47) than text (M = 4.17, SD = 0.58, $p < 0.05$). Third, self-consciousness emerged as another barrier that made students (P2, 13) concerned about the way their speech would be presented. This was especially evident among ESL students, but overall, participants felt more anxious when making voice comments (M = 3.00, SD = 1.35) than when creating text (M = 1.83, SD = 0.94). Finally, participants pointed out that voice comments were not an option when students were in a library or an office (P4), when a roommate was sleeping (P7, 11), or they were without a headset (P11).

Mid-Deployment Improvements

User feedback from the first deployment block echoed the first deployment (i.e., term paper feedback) findings in that RichReview^{++} makes the consumption of annotations very easy. On the other hand while the students in the first deployment thought that RichReview was very good for conveying higher-level comments, students in the peer discussion setting did not find the system as useful in that area. To address this problem, we decided to add two new features before running the second block: General Note and Booklet.

General Note

To encourage students to create more synthetic comments, we simply added a dedicated white page for high-level synthetic comments. The two column General Note page consists of the left column for general topics without specific anchor in the text, and the right column has the titles of each article like table of contents so that the users can talk about a meta-level topic by commenting under the title line or connecting multiple articles.

Booklet

The class was assigned several articles to be read each week, and participants found it quite difficult to connect elements from different text using Richreview^{++} -- an important aspect

in practice [6]. To address this problem, the booklet feature combines a set of multiple articles into a single session, much like a booklet. As shown in Figure 5, when a user hovers the cursor over the left-top edge of the screen, an article selection menu slide-opens, just like an interactive table of contents. A General Note was placed on top of the Booklet like a cover page.

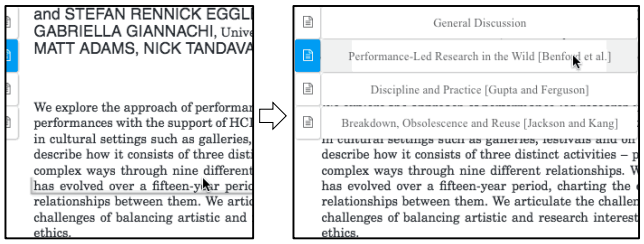

Figure 5. Booklet. Hovering the cursor over the left-top end of the screen slide-opens a cross-document navigation menu.

Result from the Follow-Up Study

In the second block of the study, participants made a similar number of textual comments (236) but much fewer voice comments (8) than in the first block. A total of 11 participants were interviewed and 13 responded to the online survey. Participants found the Booklet feature to be very useful to switch between documents (M = 4.36, SD = 0.67). Further it made peers' comments more discoverable: Because the Booklet combines the articles into a single session, the entirety of the week's comments were presented in a single bucket of Comment History. At the same time, the General Note function was barely used by the students, resulting in only 2 comments made by a single student. Lower rates of commenting might have been caused by the increasing workload toward the end of the semester.

DISCUSSION

Better than Being There?

Some of the students using RichReview believed that it was preferable to in-person meetings during the instructor's office hours. This was surprising since face-to-face meetings are typically regarded as the gold standard for communication technology. Our result suggests that our experiences with RichReview could pave a way toward educational tools that surpass the standard face-to-face meeting experience. There are two different directions that this process can take.

First, Hollan et al. noted that the goal of communication technologies is not necessarily to replicate face-to-face meetings [15]. Tools for digital collaboration can often provide unique functionality that is not available in face-to-face meetings. For example, RichReview's quick replay, asynchronous workflow, and visually obvious pointing gestures made it easier for students to thoroughly process instructor comments at their own pace. With improvement, tools like RichReview may grow to become an attractive platform that serves as an *alternative* to face-to-face meetings.

Second, a communication tool like RichReview might be used to supplement the face-to-face meeting, rather than substitute for it. For example, a new variant of RichReview could be used during a face-to-face conversation to record what is said about the document. At the end of the meeting both parties would end up with a record they can replay as they wish.

Balanced Support of Voice Annotation

One of the more interesting findings of our deployments is that participant responses focused on the difficulty of creating voice annotations, not consuming them. This effect is opposite to the typical situation in CSCW applications where speech places an unequal burden on the recipient [11]. One reason why this might be the case is because our system explicitly attempts to simplify the voice consumption task via the waveform indexing feature. Therefore, we believe this feature might generally be useful for solving the consumption problems in other applications where collaborators exchange speech data.

Our study results also point to several promising ways for enhancing voice production. Employing an automatic speech transcription would be a good way to enhance the editing capability of voice comments. As presented by Yoon et al. [38], segmenting speech into words enables word-level deletion of a part of a voice-recorded comment. When high quality auto transcription becomes available (e.g., via crowd-based systems [18]), editing voice as if it is text might become an option [28].

Overcoming the situational and social barriers of voice production is another important issue to resolve. Although some of the discomfort of creating speech annotations will fade away as they become more widespread, self-consciousness when creating voice comments might be far more difficult to overcome. One solution may be to employ a combination of speech recognition and synthesis to create comments containing emotional inflection while removing aspects of speech that make people feel self-conscious. This feature would also make it possible to use RichReview in context requiring anonymity such as scientific reviews.

Implications for Instructional Design

Even though the goal of our study was to draw out lessons for the design of new educational collaboration tools, our results also have implications concerning the *instructional design* of courses that employing these new tools. There are two major considerations that emerged.

First, instructors must understand the pedagogical affordances of different tools and chose the one suitable for one's learning goal. For example, maintaining a manageable chunk size for RichReview^{++}'s comments encourages easier peer content production and consumption. This implies that the tool can be subtly altered to promote formative and generative learning processes that values diverse perspectives.

Second, participants in the peer discussion deployment tended to make less cohesive comments even when provided with the General Note feature because it remained easier to create detail oriented comments. This suggests that appropriate incentives should be designed into the lesson plan to motivate students to take a desired action in classroom. Future systems might incorporate a discussion dashboard feature that summarizes the number of comments made by students or discussion goal and achievement status.

CONCLUSION

In this paper we conducted two deployment studies to gauge the suitability and efficacy of a multi-modal document annotation system for supporting various activities encountered in university classrooms. Our studies examined the use of the RichReview annotation tool as a mechanism through which instructors to provide feedback and students can engage in discussion about reading assignments.

We found that multimodal commentaries can successfully deliver instructor's high-level ideas, nuances, and emotions, and that students can easily understand the recorded comments by using the quick voice indexing feature. In some cases, students preferred the multi-modal comments over a face-to-face meeting with the instructor. We discovered in the peer discussion activity that the commenting method employed influenced the scope at which discussions occurred. We also found that students found text and audio comments anchored in context with the material they refer to be useful and compelling. Lastly, our results highlighted some open issues regarding the creation of speech-based annotations in this domain.

Although we found that multi-modal annotations that RichReview^{++} offers to be useful, it is important to note that the technology on its own offers only a partial solution. We believe that understanding affordances and implications of this type of collaboration tool and using it as part of a broader effort in the classroom is necessary to realize the full potential of the tool.

REFERENCES

1. Alrushiedat, N. and Olfman, L.Anchored Asynchronous Online Discussions: Facilitating Participation and Engagement in a Blended Environment. *HICSS*, 2012, 11–20.
2. Anderson, R., Davis, P., Linnell, N., and Prince, C.Classroom presenter: Enhancing interactive education with digital ink. *Computer*, (2007).
3. Anderson, R.Classroom Presenter: a classroom interaction system for active and collaborative learning. *WIPTE*, (2006).
4. Brush, A.J.B., Bargeron, D., Grudin, J., Borning, A., and Gupta, A.Supporting interaction outside of class: anchored discussions vs. discussion boards. (2002), 425–434.
5. Chalfonte, B.L., Fish, R.S., and Kraut, R.E.Expressive richness. *SIGCHI*, ACM Press (1991), 21–26.
6. Chen, N., Guimbretiere, F., and Sellen, A.Designing a Multi-slate Reading Environment to Support Active Reading Activities. *ACM Trans. Comput.-Hum. Interact. 19*, 3 (2012), 18:1–18:35.
7. Churchill, E.F., Trevor, J., Bly, S., Nelson, L., and Cubranic, D.Anchored conversations. *SIGCHI*, ACM Press (2000), 454–461.
8. Crook, A., Mauchline, A., Maw, S., et al.The use of video technology for providing feedback to students: Can it enhance the feedback experience for staff and students? *Computers & Education 58*, 1 (2012), 386–396.
9. Dorn, B., Schroeder, L., and Stankiewicz, A.Piloting TrACE: Exploring Spatiotemporal Anchored Collaboration in Asynchronous Learning. (2014).
10. Gao, F.A case study of using a social annotation tool to support collaboratively learning. *The Internet and Higher Education*, (2013).
11. Grudin, J.Why CSCW applications fail: problems in the design and evaluationof organizational interfaces. *CSCW*, ACM Press (1988), 85–93.
12. Gutwin, C., Stark, G., and Greenberg, S.Support for workspace awareness in educational groupware. *CSCL*, Association for Computational Linguistics (1995), 147–156.
13. Guzdial, M. and Turns, J.Effective Discussion Through a Computer-Mediated Anchored Forum. *Journal of the Learning Sciences 9*, 4 (2000), 437–469.
14. Hardock, G., Kurtenbach, G., and Buxton, W.A marking based interface for collaborative writing. *UIST*, ACM Press (1993), 259–266.
15. Hollan, J. and Stornetta, S.Beyond Being There. *SIGCHI*, ACM (1992), 119–125.
16. Kim, J., Glassman, E.L., Monroy-Hernández, A., and Morris, M.R.RIMES: Embedding Interactive Multimedia Exercises in Lecture Videos. *SIGCHI*, ACM (2015), 1535–1544.
17. Kulkarni, C.E., Bernstein, M.S., and Klemmer, S.R.PeerStudio: Rapid Peer Feedback Emphasizes Revision and Improves Performance. *Learning @ Scale*, ACM (2015), 75–84.
18. Lasecki, W., Miller, C., Sadilek, A., et al.Real-time Captioning by Groups of Non-experts. *UIST*, ACM (2012), 23–34.
19. Levine, S. and Ehrlich, S.The Freestyle System. *Human-Machine Interactive Systems*, (1991).
20. Marshall, C.C. and Brush, A.J.B.Exploring the relationship between personal and public annotations. 2004, 349–357.
21. Neuwirth, C.M., Chandhok, R., Charney, D., Wojahn, P., and Kim, L.Distributed collaborative writing. *SIGCHI*, ACM Press (1994), 51–57.

22. Nicol, D.J. and Boyle, J.T.Peer Instruction versus Class-wide Discussion in Large Classes: A comparison of two interaction methods in the wired classroom. *Studies in Higher Education 28*, 4 (2003), 457–473.
23. Nortcliffe, A. and Middleton, A.Effective assignment feedback through timely and personal digital audio engagement. *IGI Global, Hershey*, (2009), 409–428.
24. Oomen-Early, J., Bold, M., Wiginton, K.L., Gallien, T.L., and Anderson, N.Using asynchronous audio communication (AAC) in the online classroom: A comparative study. *Journal of Online Learning and Teaching 4*, 3 (2008), 267–276.
25. Pearson, J., Robinson, S., and Jones, M.PaperChains: Dynamic Sketch+Voice Annotations. *CSCW*, ACM (2015), 383–392.
26. Pol, J. Van der, Admiraal, W., and Simons, P.R.-J.The affordance of anchored discussion for the collaborative processing of academic texts. *International Journal of Computer-Supported Collaborative Learning 1*, 3 (2006), 339–357.
27. Randles, B., Yoon, D., Cheatle, A., Jung, M., and Guimbretiere, F.Supporting Face-to-Face Like Communication Modalities for Asynchronous Assignment Feedback in Math Education. *Learning @ Scale*, ACM (2015), 321–326.
28. Rubin, S., Berthouzoz, F., Mysore, G.J., Li, W., and Agrawala, M.Content-based tools for editing audio stories. *Proceedings of the 26th annual ACM symposium on User interface software and technology - UIST '13*, ACM Press (2013), 113–122.
29. Schilit, B.N., Golovchinsky, G., and Price, M.N.Beyond paper: supporting active reading with free form digital ink annotations. *SIGCHI*, ACM Press/Addison-Wesley Publishing Co. (1998), 249–256.
30. Shute, V.J.Focus on formative feedback. *Review of educational research 78*, 1 (2008), 153–189.
31. Silva, M.L.Camtasia in the Classroom: Student Attitudes and Preferences for Video Commentary or Microsoft Word Comments During the Revision Process. *Computers and Composition 29*, 1 (2012), 1–22.
32. Smith, M.K., Wood, W.B., Adams, W.K., et al.Why peer discussion improves student performance on in-class concept questions. *Science (New York, N.Y.) 323*, (2009), 122–124.
33. Stifelman, L., Arons, B., and Schmandt, C.The audio notebook: paper and pen interaction with structured speech. *SIGCHI*, (2001), 182–189.
34. Tsang, M. and Fitzmaurice, G.Boom chameleon: simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display. *UIST*, (2002), 111–120.
35. Whittaker, S., Hyland, P., and Wiley, M.FILOCHAT. *SIGCHI*, ACM Press (1994), 271–277.
36. Wilcox, L.D., Schilit, B.N., and Sawhney, N.Dynomite. *SIGCHI*, ACM Press (1997), 186–193.
37. Wolsey, T.Efficacy of Instructor Feedback on Written Work in an Online Program. *International Journal on E-Learning 7*, 2 (2008), 311–329.
38. Yoon, D., Chen, N., Guimbretière, F., and Sellen, A.RichReview: Blending Ink, Speech, and Gesture to Support Collaborative Document Review. *UIST*, ACM (2014), 481–490.
39. Zyto, S., Karger, D., Ackerman, M., and Mahajan, S.Successful classroom deployment of a social document annotation system. *SIGCHI*, ACM Press (2012), 1883.