In Case You Missed It: Benefits of Attendee-Shared Annotations for Non-Attendees of Remote Meetings

Mukesh Nathan¹, Mercan Topkara², Jennifer Lai², Shimei Pan², Steven Wood², Jeff Boston², Loren Terveen¹

¹University of Minnesota
Minneapolis, MN, USA
{mukesh, terveen}@cs.umn.edu

²IBM T.J. Watson Research
Hawthorne, NY, USA
{mtopkara, jlai, shimei, woodsp, daddyb}@us.ibm.com

ABSTRACT

Corporate meetings are increasingly being held remotely using web technologies. With such remote meetings being recorded and made available after the fact, there is a pressing need for tools to access and utilize these recordings efficiently. Our work explores the utility of using annotations generated by meeting attendees to meet this need. We conducted a controlled lab study to evaluate the benefits of sharing annotations. Attendee-created annotations were shared with non-attendees to assist them on typical information retrieval tasks. Results indicate that (a) non-attendees given access to shared annotations performed about as well as attendees provided with their own and shared annotations, (b) non-attendees were more confident in their responses when they used shared annotations as access cues into the recording than when they directly skimmed the video, and (c) attendees utilized shared annotations more than their own, with similar success and confidence as using their own annotations.

Author Keywords

web-conferencing, video, indexing, meetings, recordings, annotations, tags, enterprise, collaboration

ACM Classification Keywords

H.5.3. Group & Organizational Interfaces: Computer Supported Cooperative Work

General Terms

Experimentation, Design

INTRODUCTION

Meetings are an important venue within the enterprise for information exchange, where the capture of important points, decisions and/or actions is often done informally and manually. With the rise in geographically dispersed teams, organizations are increasingly adopting technology to hold meetings remotely over the web. Commercial webconferencing products such as GoToMeeting and Webex [13] are being widely used by today’s global companies to reduce travel and time costs. Many of these systems facilitate the automatic capture of information by recording remotely held meetings in their entirety and making them available for later viewing. Such systems can support people who did not attend the meeting but hope to catch up on what they missed. They create the possibility of using meeting recordings as archives of information.

As repositories of such recorded meetings grow, the value and utility of these stores will depend on providing tools that help users to quickly browse, find and retrieve elements of interest. Given the high time cost of viewing a recorded meeting in its entirety, there is a need for tools that assist efficient information retrieval. This is particularly true for people who missed a meeting, who frequently choose to learn about the proceedings from a colleague [1], rather than invest the time viewing its recording.

The capture of meetings has received a lot of attention in the research community over the last few decades. However, the majority of work has focused on supporting remote meetings by prototyping systems for the capture and access of these meetings [6, 19]. Many such systems not only record the meeting proceedings, but also capture associated artifacts such as user annotations made by attendees. Our work breaks new ground by being the first to systematically evaluate whether annotations created and shared by meeting attendees can assist non-attendees on information retrieval tasks. The Dogear system [14] showed how users could benefit from webpage bookmarks shared by others, creating a multiplicative effect of enterprise sharing. We examine whether people who miss meetings can benefit from using annotations shared by attendees as a tool to access video recordings of meetings.

Scope

Our long-term goal is to support distributed enterprise meetings – where attendees use a web conference along with a telephone bridge – such that attendees and non-attendees alike can easily discover and retrieve information from the recordings of meetings. As a first step, we focus on helping non-attendees retrieve information from a video recording of the meeting. We study a 1-to-N type meeting (in which a speaker has a discussion with a group) that is hosted in a meeting capture system deployed in a large enterprise. Those in attendance create time-stamped textual annotations during the meeting. These annotations are captured, synchronized with video and shared with those who missed the meeting as a possible aid for retrieving meeting-related information.
We present the results of a controlled lab study with 40 knowledge workers who used shared annotations to answer questions about the meeting after it had taken place. The study examined the following research questions:

**RQ1. Do annotations created and shared by attendees help non-attendees retrieve information from recordings?**

On the information retrieval tasks, we found that non-attendees with access to shared annotations performed (a) better than non-attendees without shared annotations, (b) about as well as attendees given shared annotations and their own annotations, and (c) about as well as attendees given only their own annotations.

**RQ2. To what extent are shared annotations utilized?**

We found that non-attendees (a) relied more on shared annotations as cues into the recording than skimming the video alone, (b) were about as successful using shared annotations or video skimming, and (c) were more confident in their answers when using shared annotations. For the same retrieval tasks, we found that attendees (a) utilized shared annotations more than their own, (b) were about as successful using shared or their own annotations, and (c) were about as confident in their answers using shared annotations or other retrieval methods.

The rest of the paper is organized as follows. We first present prior work related to meeting capture and review systems. We describe one such system using which our study was carried out. We then describe the design of a lab study to evaluate our research questions. Finally, we present the results of this study, discuss limitations of the study and conclude with a brief summary.

**RELATED WORK**

**Meeting Capture and Review Systems**

Creating systems to capture meeting-related information has received much attention over the past few decades. Such systems record meeting artifacts such as audio, video, slide images etc. and make them available post-hoc using a playback interface. Jabber [10], LiteMinutes [2], NotePals [3], and FiloChat [16] are just some systems from this rich line of research. Our current work was carried out in one such system (refer to ‘Collaborative Recorded Meeting System’ section). Unlike these systems, ours is a web-based collaborative service that focuses on supporting geographically distributed users to attend remote meetings and to browse meeting artifacts. Among past research systems, TeamSpace [6] is closest to ours. Our research focus is not on building such systems, but examining their ability to address meeting-specific retrieval needs of those who miss meetings.

**Indexing Meeting Recordings**

Prior research on meeting capture systems has highlighted two types of indices to support random access into recordings: (1) **Derived indices**, which are automatically calculated by the system, such as media stream analysis to detect slide transitions or sensors to detect who is in the room [4, 6, 10], and (2) **Explicit indices**, which are created by users, including digital bookmarks and personal annotations [8, 11]. Such explicit indices could be kept private to the creator, or shared with other meeting attendees as well as non-attendees. While users can create explicit indices as part of a regular meeting interaction process, the creation of derived indices often requires custom hardware or post-processing. Our current work builds on this body of research by evaluating the benefits of **explicit indices** to assist users in retrieving information discussed in meetings they missed.

**Recording User Annotations**

Our research builds on prior systems that co-index meeting recordings and user-created annotations [3]. In such systems, time-stamped annotations from different meeting attendees are combined and merged with a separately captured audio/video stream, thus making both the annotations and the media stream available for later playback. The Live Conference Dashboard (LCD) system [11] supports “live-tagging” during audio-only remote meetings in an enterprise setting. In a study, users preferred one-click tagging to other methods that allowed the tagging of meeting segments with multiple clicks (i.e. requiring both a start and an end point to the tag). We use this result in our work by allowing users to add tags with one-click. The LiteMinutes [2] system compiles meeting notes from different users into a single document that is correlated with the meeting recording. However, these studies mostly focus on implementation details of creating annotations during a meeting and their later access, without much empirical examination of the affordances they provide. Our work looks at advantages of having access to group annotations related to meeting recordings for retrieval tasks.

**Retrieving Meeting Information Using Annotations**

Kazman et al. evaluated the utility of keywords automatically extracted from a text transcript as **derived indices** into meeting recordings [10]. They found that having access to such keywords substantially improved the ability of meeting attendees to retrieve information from video recordings. Similarly, Banerjee et al. looked at the impact of automatic annotations **derived from** meeting structure on information retrieval [1]. In their user study, participants took less time to recall details of the meeting they attended when given annotations than when not given annotations. Our work complements this research by focusing on user annotations as **explicit indices** to assist non-attendees. We also examine the extent to which participants utilize shared annotations compared to other retrieval modalities.

The field of education is rich in studies that examine the benefits of having one’s annotations available for review during retrieval tasks. Much of this research can also be applied to meetings, if we consider lectures to be a meeting of type 1-to-N. For example, a study of ChittyChatty [7], a system that supports the capture and access of notes during a lecture, showed that students were able to accurately recall information based on their personal notes indexed to the lecture audio. A review of the field noted that three-fourths of such studies found positive effects of having access to one’s own notes [7]. Kalnikaité and Whittaker showed a correlation of shared annotations with learning [8]. Students in a longitudinal study preferred accessing previously accessed annotations while re-
Figure 1. The Live-Annotations interface allows remote meeting participants to create annotations in real-time.

visiting lecture recordings. Our work extends this body of studies by reporting on a controlled lab study focused on examining the causal relation between shared annotations and performance on information retrieval tasks.

THE COLLABORATIVE RECORDED MEETING SYSTEM
The Collaborative Recorded Meeting (CRM) system is a research prototype that allows users to host and join web meetings, record them, and view previously recorded meetings in the form of videos. The CRM system has two primary modes of use: 1) during meeting interaction, where users can create text annotations while the meeting is being recorded, and 2) post-meeting interaction, where users can view the recording as well as annotations. The features relevant to this paper are the interface that allows for during-meeting annotations to be created (Figure 1) and the ability for users to view these annotations during playback (Figure 2). The remainder of the system is not relevant to this paper and can be seen in [15].

During a meeting users can create annotations that are time-tamped and displayed at the bottom of the interface in a horizontal chronologically ordered list (Figure 1) as they are created. Only the users’ own annotations are visible to them during the meeting. Users can mark annotations as private or public (shared), the default being public. Annotations shared by other attendees are not displayed during the meeting, but can be viewed after the meeting during playback. While one could imagine benefits in broadcasting all shared annotations in real-time as they are created, we decided for this initial version not to display them due to the distraction of seeing them come up while one tries to follow the meeting [11].

The CRM system automatically aligns user annotations on the meeting timeline and displays them in a side panel next to the video player (Figure 2). Clicking an annotation on the left causes the player to automatically seek to that point in the video and start playing from there.

Figure 2. The Recorded-Meeting Playback interface. Clicking an annotation on the left starts playback from the point in the meeting video where it was created.

USER STUDY
We carried out a controlled lab study (N=40) to evaluate the impact of annotations on retrieving information from a meeting recording.

Study Design
Users access meeting records as either (a) meeting attendees who later want to retrieve details based on information partially remembered, or (b) as non-attendees who are trying to discover information about a missed meeting [17]. Our study consisted of a between-subject, 2 by 2 design (see Table 1) where attendees and non-attendees were divided into 2 conditions, which differed by whether or not they had access to shared annotations during information retrieval tasks.

Participants
Forty people (32 males and 8 females) at a large corporate research lab responded to an email invitation. They were researchers or summer interns with a background in Computer Science. All the participants were fluent English speakers, and a majority (62%) regularly participated in remote meetings. Twenty participants were assigned to each of the two types of users (attendees or non-attendees) for a total of 40 participants. Within each of these user groups, participants were randomly assigned to either the Shared or NoShared conditions (see Table 1). They were given a lunch certificate for their participation.

Study Conditions
Participants in the Attended-NoShared condition only had access to their own annotations for retrieving meeting-related information. Those in the Attended-Shared condition had access to their own annotations, plus those created by 3 other attendees, during the retrieval task. These 3 users were randomly selected from the set of participants who created annotations. In a pilot study, users found that having access to 500+ annotations from all 20 meeting attendees was overwhelming. Scanning duplicate content in annotations was
also frustrating. We observed that users could handle approximately 100 annotations comfortably. This led us to randomly pick 3 users (avg. of 27 annotations per user) whose annotations were shared. These 3 users were then assigned to the Attended-NoShared condition.

Participants in the Missed-Shared treatment were provided the annotations of the same 3 randomly selected users as Attended-Shared participants. Those in the Missed-NoShared treatment were not provided with any annotations during the information retrieval tasks.

**Procedure**

Wilcox et al [18] studied annotation capture and review behavior and found that information from previous meetings is often reviewed prior to the next meeting in order to revisit important details and refresh one’s memory. In line with this finding, we designed the study with two phases: 1) a participation phase and 2) an information retrieval phase, spaced a week apart.

1) **Participation Phase:**

During this phase, only the 20 users assigned to the Attended conditions were asked to participate in a remote meeting. They were asked to make annotations using the Live-Annotations interface (Figure 1) during the meeting. Before starting the experiment, there was a short familiarization session where users created, edited and deleted annotations using the Live-Annotations interface. They were told that annotations could be shared, or not, with other meetings by marking them as ‘private’ or ‘public’ (the default was public). They were informed that while shared annotations are not broadcast in real-time with other attendees, they might be available to others after the meeting. Next, participants attended a meeting by viewing a video stream from a remote location. In order to ensure that all 20 participants in the Attendee conditions experienced the same meeting, the video stream was actually a recording. While they were told to assume the meeting was happening live, there was no telephone conference and they could not speak or ask questions.

The recording used as the remotely streamed meeting was chosen from a corporate database of real meeting recordings. In the selected recording, a researcher presents findings from experiments related to novel text summarization techniques. The talk is a traditional 1-to-N type meeting with a presenter speaking and using a slide deck. It is representative of a common meeting style in large corporations (e.g., technical talks, recurring meetings with an invited speaker). The meeting could be easily understood by users, as verified in a post-task survey, and did not require any additional context or knowledge in order to understand the topic being presented. The talk was edited down into a logical unit (introduction and results from 2 experiments) with duration of 25 minutes. None of the users had viewed the meeting recording previously, nor were familiar with the contents of the recording. The Participation Phase took less than an hour per user.

2) **Information Retrieval Phase:**

In this phase, all 40 participants were given a series of questions to answer based on information in the meeting recording. Participants in the Missed conditions were told that they had ‘missed’ the meeting but needed to attempt to answer the questions as best they could with the information available from the recording of the meeting. For users in the Attended conditions, the Information Retrieval Phase was conducted seven days after they attended the meeting. Participants in Attended-Shared, Attended-NoShared and Missed-Shared conditions were first given a tutorial to familiarize themselves with the Recorded-Meeting Playback interface (Figure 2). They practiced the use of random access into sections of the recording based on the selection of a specific annotation. Users were also showed how to search through the provided annotations. When they were familiar with the interface, they began the experiment proper. For users in the Missed-NoShared condition, no annotations were displayed.

Prior research has shown that users extract two main types of information from meeting recordings: gist or meeting overview, and specific facts [17]. Of the two types of information (gist vs. facts), we decided to focus on factual information since it can be more objectively evaluated. We asked all participants to answer 10 factual questions. The questions required them to extract facts from the meeting, as opposed to summarizing information distributed across the meeting. All questions required paying attention to the audio and video in the recording in order to answer them correctly. Questions were framed prior to the start of the study and were independent of annotations created in the Participation Phase. A pilot study was conducted to ensure the questions could be easily understood and were of comparable difficulty. The questions were not in a multiple choice format to prevent users from randomly guessing. An example question was: “The speaker demonstrates a software tool. Briefly describe what the tool allows users to do.” The 10 questions were presented in random order. Participants were asked to answer them as quickly and accurately as possible. We imposed an upper bound of 30 minutes for the task. The Information Retrieval Phase took approximately 45 minutes per user.

**Variables and Measures**

The two independent variables in the study were Meeting Attendance {Attended, Missed} and Access to Shared Annotations {Shared, NoShared}.

The following measures were used:

**Score:** Two judges, blind to our research questions, created a key by viewing the meeting recording and generating target answers. The judges then combined their answers and

<table>
<thead>
<tr>
<th>Attendees</th>
<th>No Access to Shared Annotations</th>
<th>Had Access to Shared Annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attended-NoShared (own annotations)</td>
<td>N=10</td>
<td>Attended-Shared (own annotations)</td>
</tr>
<tr>
<td>Non-attendees</td>
<td>Missed-NoShared</td>
<td>Missed-Shared</td>
</tr>
<tr>
<td>N=10</td>
<td>N=10</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. The conditions in the 2x2 between-subjects design
specified keywords that needed to be present in each answer. Scores for each question ranged from 0 to 1 (in quarter point increments) depending on how much of the target answer participants mentioned. The judges scored the answers independently with disagreements being resolved by a third judge.

**Time on Task:** The question-answering interface automatically recorded the total time participants took to answer all questions, starting from the time they viewed the first question until they submitted an answer to the last question.

**Retrieval Method:** After each question, participants were asked to note the exact method they employed to answer the question. The options provided were: (1) **User Annotation** – Participants with access to annotations were instructed to indicate this option when the content of an annotation pointed them to a relevant segment in the video or in some cases, directly contained the answer\(^1\). Further, participants in the Attended-ShARED condition also noted when they used their own annotations as opposed to a shared annotation. (2) **Video only** – In cases where participants answered questions purely by skimming through the video, they were instructed to select this option. (3) **Memory only** – This option was selected when users relied purely on their own memory, without having to use aids such as the video or annotations.

**User Confidence:** Participants indicated their confidence for each of their answers. A 3-point Likert scale with values ‘very confident’, ‘somewhat confident’ and ‘not at all confident’ was used.

---

\(^{1}\)No distinction was made between using annotations only and using annotations to access video. The former occurred very infrequently. Users referred to the video even in cases where the annotation contained the answer. Post-task questioning revealed that users wanted to verify the annotation content by viewing the video.

---

**CONTENT ANALYSIS OF LIVE MEETING ANNOTATIONS**

During the Participation Phase of the study, 522 annotations were created by 20 users who were assigned to the two Attended conditions. On average, users entered 6 words per annotation. As noted by Kelkar et al. [11], annotations created during a meeting tend to be lengthy when compared to tags created on static data on the web. 30% of the annotations had more than 10 words, and 16 users added at least one annotation that had more than 10 words (e.g. “typical behaviors: with word excision, people are getting through the text quicker”). However, the majority of the annotations were not long. With short annotations, users created pointers to important parts of the meeting with keywords such as “fish eye”, “graph”, and “architecture”. 48% of the annotations have less than 5 words. Here are some examples: “cool graphs”, “user prefs”, “demo time!”, “excision - people miss stuff”, “this technique is old”.

The short annotations are very similar to the type of tags commonly seen in web tagging systems and are used as placeholders for rediscovering content. It seemed users in our study were creating ‘bookmarks’ into the meeting that could then be used to retrieve details if required. As one participant noted: “I like to capture as much information in my tags so I don’t need to view much of the video later”. By contrast, notes created during a lecture are different in nature as they typically summarize or clarify the content being recorded, list questions and objections in mind, etc. [9]. In a survey, we asked participants to indicate whether their annotation behavior was influenced by the knowledge that the meeting recordings could be played back at the exact points where annotations were made. Participants strongly agreed with this view, providing credence to our notion of users being ‘bookmarkers’.

We found the same content repeated in many annotations (e.g. 11 different users created an annotation with the term “fish-eye”). Such repetition could be used to create a cluster or topic flow graph of the meeting. Repetition might be reduced if annotations were broadcast live among users as they are created, but as mentioned earlier we decided against it due to the additional cognitive load from the display of all shared annotations in real-time [11].

Another phenomenon we observed was that when several key concepts were mentioned back-to-back in the meeting, users were not able to capture all of them in their annotations. For example, at one point during the meeting the speaker introduces the term “strategic information processing”, and within 1.5 seconds he introduces the concept of “golden triangle”. Five users annotated “golden triangle”, and three others recorded “strategic information processing”. None of the users noted both in their annotations. This highlights the potential value of sharing during-meeting annotations where the speaker is not presenting in a structured educational lecture format that usually provides room for the students to digest terms that are being introduced.

The duration for entering free-form text annotations was on the average 17 seconds (max = 99 secs). We also observed that on average, users left 56 seconds between annotations and the minimum break was 1.2 seconds (max = 9 mins). These two data points suggest that it is important to provide users the ability to add quick, low-effort annotations during a meeting that will assist them in reviewing the details later.

**RESULTS**

We now discuss our findings from the Information Retrieval Phase of our study that evaluated participants answering factual questions based on the meeting. User performance on these tasks is summarized in Table 2.

Results of the retrieval tasks were first analyzed using a two-way ANOVA with Score as dependent variable, and Attendance {Attended, Missed} and Access to Shared Annotations {Shared, NoShared} as independent variables. A significant main effect for Access to Shared Annotations (F\(_{1,36}\)=16.66, \(p<0.001\)) was observed, indicating that participants with access to shared annotations scored higher, independent of meeting attendance.

There was also a main effect for Attendance, suggesting that participants performed better on the retrieval tasks when they attended the meeting (F\(_{1,36}\)=6, \(p<0.05\)). The interaction effect for Shared Annotations and Attendance was not significant, (\(p>0.1\)). A similar ANOVA with Time on Task as dependent variable found no significant results, (\(p>0.1\)).
ing tested the main effects, we proceeded to answer our research questions.

**RQ1. Do annotations created and shared by attendees help non-attendees retrieve information from recordings?**

We wanted to clearly understand the value annotations created by meeting attendees could provide people who missed the meeting. For this, we compared the **Missed-Shared** condition to the 3 other study conditions. (1) First, the performance of **Missed-Shared** users was compared with that of participants in the **Missed-NoShared** condition. This contrast served as a baseline. (2) Next, the **Missed-Shared** and **Attended-Shared** conditions were compared. This examined the advantage meeting attendance had during the retrieval tasks, keeping access to shared annotations constant. We wanted to see the extent to which providing annotations as retrieval aids could help compensate for having missed the meeting. (3) Finally, we wanted to see if access to annotations created by many users could provide non-attendees an edge over attendees who had access only to their own annotations. We explored this case by comparing the **Missed-Shared** and **Attended-NoShared** conditions.

Since there was a significant effect for **Access to Shared Annotations** on the **Score** for retrieval tasks, and not for **Time on Task**, we only compared the different conditions based on scores. The pair-wise comparisons were made using Tukey’s HSD. We now discuss the results of the 3 comparisons as summarized in Table 3.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Score [%]</th>
<th>Time [mins:secs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attended-NoShared</td>
<td>72 (±16)</td>
<td>22:05 (±4 : 43)</td>
</tr>
<tr>
<td>Attended-Shared</td>
<td>87 (±8)</td>
<td>22:23 (±4 : 42)</td>
</tr>
<tr>
<td>Missed-NoShared</td>
<td>63 (±9)</td>
<td>27:04 (±1 : 27)</td>
</tr>
<tr>
<td>Missed-Shared</td>
<td>78 (±11)</td>
<td>22:15 (±4 : 12)</td>
</tr>
</tbody>
</table>

Table 2. The performance (mean score and time on task) of users on retrieval tasks, along with standard deviations.

(1) **Missed-Shared vs. Missed-NoShared**

**H1: Missed-Shared users will perform better than the Missed-NoShared users on questions about the meeting.**

Shared annotations act as semantic indices into the meeting recording and help users to quickly locate key ideas.

Participants in the **Missed-NoShared** condition had no access to any annotations, while those in the **Missed-Shared** condition were provided with 67 annotations shared by attendees\(^2\). The results supported our prediction, with **Missed-Shared** participants performing significantly better than **Missed-NoShared** users. **Missed-Shared** users seemed to benefit on accuracy by having access to annotations made by attendees. They were able to effectively utilize shared annotations to answer the questions, while **Missed-NoShared** users had to rely only on skimming through the video to understand the structure of the meeting.

\(^2\)Shared annotations directly contained answers to 2 of the 10 questions. In 6 cases, they had keywords pointing to the general region in the recording that contained the answer. It was up to participants to recognize and utilize them.

(2) **Missed-Shared vs. Attended-Shared**

**H2: Attended-Shared users will perform better than Missed-Shared users; having attended the meeting, their own memory and personal annotations will prove useful for information retrieval purposes.**

The result that participants in the **Missed-Shared** condition performed better than those in **Missed-NoShared**, though seemingly intuitive, is non-trivial. They did so by relying heavily on shared annotations. Though attendees shared almost all annotations they created, this was more to do with the default setting and not an explicit desire to share. Contrary to past findings [12], our informal interviews with **Attendee** users revealed that they did not change their annotation content despite knowing their annotations may be shared. This meant that annotations were primarily created for the benefit of their author. The fact that the **Missed-Shared** users could utilize potentially esoteric shared annotations makes this result interesting. It highlights the value of sharing annotations: non-attendees can exploit annotations made by someone else to navigate to important parts of meeting recordings.

Table 3. Pair-wise comparisons of scores using Tukey’s test. The lower and upper confidence limits are given. Difference in scores between the compared conditions along with the adjusted P values are shown. Significant differences are indicated using *.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>diff</th>
<th>lower</th>
<th>upper</th>
<th>P adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missed-Shared vs. Missed-NoShared</td>
<td>1.45</td>
<td>0.05</td>
<td>2.85</td>
<td>0.04*</td>
</tr>
<tr>
<td>Missed-Shared vs. Attended-Shared</td>
<td>-0.95</td>
<td>-2.35</td>
<td>0.45</td>
<td>0.28</td>
</tr>
<tr>
<td>Missed-Shared vs. Attended-NoShared</td>
<td>0.60</td>
<td>-0.80</td>
<td>2.0</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 3. Pair-wise comparisons of scores using Tukey’s test. The lower and upper confidence limits are given. Difference in scores between the compared conditions along with the adjusted P values are shown. Significant differences are indicated using *.

Contrary to our prediction, we were surprised to see that there was no significant difference in scores for these treatments (see Table 3, row 2). Access to shared annotations alone seemed to be sufficient to perform comparably well on the retrieval tasks, irrespective of attendance. For **Attended-Shared** users, having their own personal annotations as well as memories of the meeting did not provide a statistically significant advantage for the purpose of retrieving factual information from the meeting recording. Having participated in the meeting, attendees were able to remember its general structure, as they indicated in the post-task interviews. However, the questions in the Information Retrieval Phase required extracting specific details from the audio/video recording, details that could have been hard to recall from memory. As one participant commented: “For the most part I had forgotten details in the video from when I attended the first part of experiment. My comments reminded me of what I had seen.” We must point out that **Attended-Shared** users did not perform poorly. They did exceptionally well as shown by the high success rate (87%). It is just that non-attendees given shared annotations performed comparably well (78%).
(3) **Missed-Shared vs. Attended-NoShared**

**H3:** Access to one’s personalized annotations will help Attended-NoShared users outperform Missed-Shared users. Reasons mentioned in hypothesis H2 also apply.

We found no significant difference between the scores of these two treatments, (see row 3 in Table 3). One would expect that participants who had attended the meeting earlier would have an advantage over those who missed the meeting. This was not the case in our study. On the other hand, the absolute scores for those who missed the meeting were better than the attendees (see Table 2).

Once again, we were surprised to see that Missed-Shared users performed nearly as well as participants who had attended the meeting. The Missed-Shared users had access to 67 shared annotations while Attended-NoShared only had their own (avg=27) annotations for assistance on retrieval tasks, besides their organic memory. This seemed to have leveled the playing field on retrieval tasks, as shared annotations seemed to help Missed-Shared participants to compensate for not having attended the meeting. We speculate that this could also be because of the factual nature of our questions. Attendees would possibly perform better on gist type questions that require an overall understanding of the meeting content. Nevertheless, these findings were intriguing and require further exploration in the future.

Having established that participants with access to shared annotations performed well on the retrieval tasks, despite having missed the meeting, we looked more closely at the groups that had access to shared annotations – Attended-Shared and Missed-Shared – to examine the extent to which shared annotations were actually used. What role did shared annotations play in the success of Missed-Shared participants? Why didn’t personal annotations give Attended-Shared users an advantage over Missed-Shared users? We examine such questions in RQ2.

**RQ2. To what extent are shared annotations utilized compared to other available retrieval methods? Can using shared annotations lead to more successful outcomes?**

During the Information Retrieval Phase, participants were asked to indicate whether they used ‘video skimming’ or ‘annotations’ to arrive at the answer, after answering each question. For each question, they also indicated their level of confidence in their answer. The usage of different retrieval methods, the success of using the different methods and user confidence in their answers are shown in Table 4 for Missed-Shared users. Table 5 shows similar values for Attended-Shared users.

(1) **Missed-Shared condition**

**H4:** In the absence of personal memories from attendance or other any other aids, Missed-Shared subjects will rely more on shared annotations than video skimming to identify segments relevant to the retrieval tasks.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Annot.</th>
<th>Video</th>
<th>t(9)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Retrieval Method</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usage</td>
<td>90%</td>
<td>10%</td>
<td>10.96</td>
<td>0.00*</td>
</tr>
<tr>
<td><strong>Success Rate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>68%</td>
<td>47%</td>
<td>1.48</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Confidence in Answers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>70%</td>
<td>30%</td>
<td>2.28</td>
<td>0.04*</td>
</tr>
</tbody>
</table>

Table 4. Comparison of ‘Annotations’ and ‘Video Skimming’ alone as retrieval methods for Missed-Shared condition using paired t-tests. Success rates indicate % of answers with a score > 0.5 Confidence scores indicate the % of answers marked ‘Very Confident’ by users. Averages of usage, success and confidence are shown.

**Usage:** Our hypothesis was confirmed. Participants used shared annotations significantly more compared to video skimming to locate relevant information (see Table 4). In a survey, Missed-Shared participants were asked to rate the usefulness of having shared annotations available during the retrieval tasks on a 1-5 Likert scale. They unanimously indicated (avg. score=5) that having shared annotations helped them perform well on the tasks (“shared annotations are super-useful”) and that they would use a system with shared annotations if it was available.

We asked all participants about their general strategy to answer questions. For participants who had access to annotations, the most common approach to answering a question was to 1) ‘search’ for keywords/synonyms in the annotation content 2) ‘browse’ through the annotations if search failed 3) ‘skim’ through the video as a last option. Users tried to avoid viewing the video as far as possible. Annotations pointed to the general area in the recording where information potentially relevant to a question was. They then used their judgment to answer the question.

**Success:** There was no statistically significant difference in the success rates when participants used annotations created by someone else, or when they depended solely on their own ability to skim the video to answer a question. User feedback towards annotations was generally positive. Participants liked having access to annotations and were “surprised by how useful annotations are for finding information in the video.”

**Confidence:** Participants were more confident in their answers when they used shared annotations than when they used the video alone. Clicking on and following annotations increased their chances of getting to the correct region in the video. More often than not, the region contained the information to be retrieved, as is evidenced by the high success rate using annotations. The video and annotation content reinforce each other, thereby instilling confidence.

(2) **Attended-Shared condition**

**H5:** Meeting attendees will rely more on their own annotations than other methods to answer retrieval questions; they will understand the organization of their own annotations better. Further, we predicted that attendees would be more confident in answers using their own annotations, as these will trigger one’s memory and instill confidence during retrieval tasks.
<table>
<thead>
<tr>
<th>Measure</th>
<th>Own Annot.</th>
<th>Shared Annot.</th>
<th>Video &amp; Memory</th>
<th>F(2, 27)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usage</td>
<td>18%</td>
<td>72%</td>
<td>10%</td>
<td>62.2</td>
<td>0.00*</td>
</tr>
<tr>
<td>Success</td>
<td>58%</td>
<td>89%</td>
<td>47%</td>
<td>3.15</td>
<td>0.04*</td>
</tr>
<tr>
<td>Confidence</td>
<td>72%</td>
<td>69%</td>
<td>45%</td>
<td>1.50</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Table 5. Attended-Shared participant’s usage of annotations (own and shared), video skimming and organic memory for retrieval. Success indicates % of answers with a score > 0.5 Confidence = % of answers marked as ‘Very confident’. Each row shows results of a one-way ANOVA. Averages of usage, success and confidence are shown.

**Usage:** Our results showed a significant difference in retrieval method usage (see row 1 in Table 5). Follow up comparisons (using Tukey’s HSD) showed a significant difference between shared annotation usage and video/memory alone (p < 0.001). Contrary to our prediction, Attended-Shared participants utilized shared annotations significantly more than their own (p < 0.001).

One reason for the high usage of shared annotations could be the sheer number of them available. All Attended-Shared users had access to 67 shared annotations as opposed to an average of 27 (std. dev=12) personal annotations. The playback interface clearly separated personal and shared annotations into separate sections (see Figure 2), so participants could easily restrict themselves to using their own annotations if they wanted to.

A common sentiment was that “Finding something just using the video requires lengthy playback and that’s a problem”. This was something we hoped annotations would address by acting as indices into the video stream. When provided with both their own annotations from the meeting as well as annotations shared by other users, attendees did not favor their own annotations over those shared by others, as long as they helped reduce video playback.

On the post-task survey, participants indicated that shared annotations helped them on the retrieval tasks (avg. score =4.1 out of 5). In the case of 4 of the 10 Attended-Shared users, they rated the usefulness of their own annotations low and those of shared annotations high. These participants used shared annotations to answer 72% of the questions, and performed as well as other participants in their treatment condition (average score=85%). The utility of shared annotations was evident: “Good thing I had access to shared comments because my comments were useless!” Shared annotations made up for the inadequacies of personal notes. One user felt that sharing annotations even resulted in “division of labor where lots of people commenting around the same video can provide benefit to everyone else."

**Success:** We wanted to see if using annotations to answer questions lead to more successful outcomes for attendees. Results indicated a significant difference between success rates when using various retrieval methods (as seen in row 2 of Table 5). Post-test comparisons (using Tukey’s HSD) showed a significant difference in success rate between using shared annotations and video/memory (p < 0.05). Most importantly, there was no significant difference in success when attendees used their own annotations or those shared by others. They were about as likely to get an answer correct, irrespective of whose annotations they used.

**Confidence:** There was no significant difference in confidence using any of the 3 retrieval methods (row 3 in Table 5). When using (shared or own) annotations to answer questions, participants were equally confident in their answers irrespective of the creator of the annotation. Subjective feedback from users shed some light on this unexpected outcome. Since users did not favor their own annotations over shared annotations while searching for answers, once they found success utilizing another user’s annotations, they began to “trust shared annotations”. When shared annotation content was not clear, participants played the recording to confirm the intent of the annotation creator. This could be one reason for having the same confidence in their answers, using personal or shared annotations.

The results from RQ2 shed some light on why Missed-Shared users performed as well as Attended-Shared users, as per the results of RQ1. Attended-Shared users relied heavily on the 67 shared annotations. Their high success rate using others’ annotations shows that shared annotations greatly helped in locating answers in the meeting recording. Since the same 67 annotations were shared with Missed-Shared users, they were able to reap similar benefits as Attended-Shared users.

The value of shared versus personal annotations is further highlighted by three observations:

(a) There was a significant difference between the score of Attended-Shared and Attended-NoShared users (p < 0.05, using Tukey’s HSD).

(b) It was not always possible for attendees to capture all key events in their annotations ‘live’ during the meeting resulting in personal annotations missing details. This was seen in our Content Analysis as well as in feedback from Attended participants. Further, recent research has shown that users often expend effort recording information in annotations that may turn out to be irrelevant [7].

(c) Attendance did not seem to provide any clear advantage for the purposes of retrieving factual meeting-related information. With a one-week gap between the two study phases, utilizing annotations as memory cues could have become more challenging, as seen in [7]. Personal annotations may also turn out to be ineffective if users can no longer remember what they mean [16].

**LIMITATIONS AND FUTURE WORK**

**Interacting During Remote Meetings**

Our current study reports results from a single 1-to-N style meeting. While we believe that this type of meeting is representative of a large number of corporate remote meetings, it is a limitation of this work that users were not able to interact with the speaker or other attendees. In reality, there are a certain number of these types of presentations that are closed to question asking due to the large number of people on the call. In ‘remote’ meetings participants do not have the same level of attention as face-to-face meetings, since they are distracted by email and IM. While the setup of needing only to watch
the meeting instead of directly engaging may differ from real
life settings, we believe it sets the upper limit on annotation
activity level. Also, having the annotation UI available gave
users “something to do” and helped them “stay focused” as
they commented. Different meeting styles could generate dif-
ferent levels of annotation activity and the gains of sharing at-
tendee annotations will have to be examined across different
meeting styles.

Handling Information Overload
Based on results from a pilot study we decided to share the
public annotations of only 3 of 20 meeting attendees with oth-
ers as shared annotations since otherwise the amount of dup-
licate annotations would be overwhelming. We envision that
real world systems will use a combination of information fil-
tering and interaction techniques to prevent information over-
load. More research would be needed to understand the best
mechanism for presenting large numbers of annotation (e.g.
greater than 100 annotations).

Real-time Sharing
As implemented, our system does not broadcast annotations
as they are being generated. Few users mentioned that it
would be useful to have such realtime exchange of annota-
tions. They felt that it could inspire new annotations of their
own as well as help them avoid repeating the same anno-
tations as others. Broadcasting tweets around a shared hashtag
at events is a comparable practice that people already indulge
in today. However, viewing (redundant) shared annotations
could lead to a processing overhead. The trade-offs involved
would need to be examined more closely.

Annotation Browsing Interfaces
In many cases users had the information needed to answer
a question available in the annotations (own and/or shared),
but failed to find it. They either did not use the right key-
word in their browser-search or skimming through all the an-
notations was not easy and they gave up. This points to the
vital importance of presenting annotations in a way that fa-
vors effective browsing. For example, many shared annota-
tions repeated the same information. Common items across
all shared annotations could be extracted and made into a hy-
perlinked summary to present the important points in a meet-
ing. Another approach would be to provide a better search
interface tailored to during-meeting annotations that possibly
contain repetitious information and typos.

Explicit vs. Derived Indices
Our results showed that user-generated annotations co-
indexed with meeting recordings could be a useful informa-
tion retrieval aid, especially for non-attendees. However,
it remains to be seen how such explicit indices would fair
against machine-generated derived indices. While retrieving
verbatim facts from the recording via derived indices (e.g.
text transcript) could be accomplished, we conjecture that the
value of explicit indices will be in retrieval tasks that rely on
interpretations of information shared in meetings.

CONCLUSIONS
We present the findings from a lab study (N=40) that exam-
ined the usefulness of sharing annotations created during a
web conference meeting with those who missed it. The con-
text for our research is enterprise meeting capture and review
systems, which allow distributed knowledge workers to par-
icipate in and record meetings. These recordings are avail-
able to attendees as well as non-attendees with the need to re-
trieve specific information that might be hard to find using au-
томatically created indices for meeting recordings (e.g. slide
transitions). Our study evaluated the ability of non-attendees
to retrieve information using shared annotations by measur-
ing their performance on retrieval tasks, which consisted of
a set of factual questions about a meeting. We wanted to see
if having access to annotations could aid non-attendees to an-
swer the factual questions.

Our results indicate that having access to shared annotations
substantially improves the ability of non-attendees to retrieve
relevant information. When provided with shared annota-
tions, non-attendees performed better than those without ac-
cess to such annotations. We found that non-attendees fre-
quently used shared annotations over other retrieval methods,
and were significantly more confident in their answers when
doing so. This was true of attendee participants as well, who
utilized annotations created by others more than their own,
underlining the value of shared annotations. Most interest-
ingly, we observed that non-attendees given access to shared
annotations performed about as well as those who had at-
tended the meeting. The main difference between attendees
and non-attendees, as reported in [17] is that for attendees the
retrieval process is one of assisted reconstruction, whereas
non-attendees have no prior information available to struc-
ture their access. Our results indicate that shared annotations
can partially address this disparity by acting as retrieval cues
for non-attendees.

Shared annotations assist in augmenting individual memory,
leading to what Engelbart called “Collective IQ” where a
group can “leverage its collective memory, perception and ex-
perience into applicable knowledge”[5]. Storing partici-
pant annotations along with meeting recordings in organi-
zational multimedia repositories can help build a knowledge
base of semantic information. Such enhanced records can
then prove beneficial to users across the organization, and not
just those who participated in the meeting originally. In line
with these ideals, our work shows how attendee-generated
annotations can be valuable to non-attendees in distributed
enterprise meeting systems. Sharing of annotations between
attendees and non-attendees can address the need for rele-
vant indices into meeting records such as those found in ever
growing corporate meeting recordings repositories.

REFERENCES
1. Banerjee, S., Rose, C., and Rudnicky, A. The necessity
of a meeting recording and playback system, and the
benefit of topiclevel annotations to meeting browsing. In
Human-Computer Interaction - INTERACT 2005.


