Bringing Mobile into Meetings: Enhancing Distributed Meeting Participation on Smartwatches and Mobile Phones

Scott Carter, Jennifer Marlow, Aki Komori
FX Palo Alto Laboratory, Inc.
Palo Alto, California, USA
{carter | marlow | komori}@fxpal.com

Ville Mäkelä
University of Tampere
Tampere, Finland
Ville.Mi.Makela@uta.fi

ABSTRACT
Most teleconferencing tools treat users in distributed meetings monolithically: all participants are meant to be interconnected in more-or-less the same manner. In practice, people connect to meetings in different contexts, sometimes sitting in front of a laptop or tablet giving their full attention, but at other times mobile and concurrently involved in other tasks or as a liminal participant in a larger group meeting. In this paper, we present the design and evaluation of two applications, MixMeetWear and MixMeetMate, to help users in non-standard contexts flexibly participate in meetings.

Author Keywords
Mobile; teleconferencing; wearable

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

INTRODUCTION
Meetings involving video conferencing have long allowed distributed teams to communicate and collaborate. While many meetings take place in office rooms with desktop or laptop streaming video, as the workforce becomes more mobile, finding ways to allow mobile users to maintain awareness and participate is increasingly important. Recent market research found that nearly 75 percent of organizations intend to introduce mobile video conferencing. Of that group, 61 percent want to integrate video conferencing within a unified platform [5].

While many distributed meeting tools support users connecting to a meeting from several different devices, they tend to expect relatively symmetric user goals and contexts.

MobileHCI '16, September 06 - 09, 2016, Florence, Italy

ACM 978-1-4503-4408-1/16/09…$15.00
DOI: http://dx.doi.org/10.1145/2935334.2935355

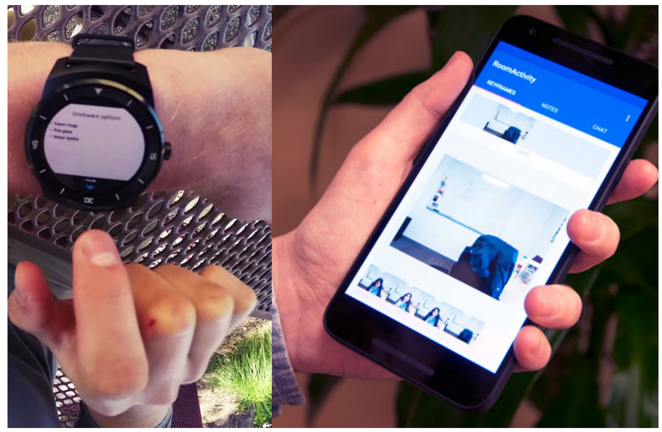


Figure 1. MixMeetWear (left) a smartwatch app for following meetings remotely, and MixMeetMate (right), a mobile app that augments collocated meetings.

For example, in a three-way meeting between a group of three people in a conference room, a remote user at a desk, and a mobile user, the first group might stream a webcam and microphone attached to a desktop, the second a laptop webcam and mic, and the third might use a phone app and the mobile device's camera and mic. However, with current distributed meeting technologies, each user would be represented more-or-less equivalently and have similar controls. Based on prior observations of many different types of meetings, we believe that differences in user contexts in meetings often require different interface designs, representations, and features.

In this work, we focus on situations in which users need to maintain awareness of a meeting while accomplishing other tasks, including searching for or capturing content related to the meeting topic. There is a continuum of situations that require divided attention in meetings; here we focus on two distinct scenarios that circumscribe the design space of requirements: users in the "hub" of a hub-and-spoke distributed meeting who contribute to the meeting through backchannels rather than as fully fledged participants, and users who are mobile, connected to the meeting via a smartwatch, and need to accomplish some other task.

RELATED WORK

Backchannel work in distributed hubs

Previous work has focused on creating backchannels within video-conferencing systems or groups of people in a room to allow participants (who are all themselves connected to the meeting) to share content and materials. Such work has addressed participation through tablet and mobile devices in a variety of different ways. For example, the SIDEBAR system focused on providing supplementary information to participants in a video conference on a tablet. This system used face-tracking to let users recognize who else was in the meeting and look up supplementary information about their profiles and locations, as well as engage in sidebar communications with a backchannel [4]. An additional tablet-based system, Tin Can, was set up to promote more equal participation in class discussions [8].

Other work has looked at ways of tying individual mobile devices together to create a unified display or sharing space. These include the Ubi-jector system [15] designed to allow participants to switch between personal and shared workspace areas, and was designed in particular for ad-hoc meetings in poorly equipped settings. Lucero et al. focused on enabling unified interaction spaces for mobile collocated users [12, 13]. Jokela et al. [11] also looked at ways of binding together different mobile devices in a group meeting with their Easy Groups and FlexiGroups tools. Finally, View & Share [7] enabled sharing photos between different mobile devices in a collocated group meeting, so not everyone had to cluster around a single (small) display.

While these works share a similar goal of allowing groups to share and view information, it should be noted that they all address the use case of a collocated group of people. Additionally, they address people with a similar level of mobility and participation levels (e.g. everyone is seated in a room and not necessarily on the go). It is likely that changing these two dimensions of distributed meetings (increased mobility and a partially distributed arrangement such as a hub and spoke) will impact social dynamics and behaviors due to the fact that not everyone shares a common environment and point of reference.

Mobile tools and distributed attention

Many off-the-shelf meeting tools also fail to integrate the context of distributed users, making it challenging for them to follow along with and contribute to meetings. In particular, current smartphone solutions do not answer needs of users who are engaged in other tasks, but would still like to stay aware of a meeting. In this context, smartphones tend be disruptive, as they are not glanceable unless specifically held up, which presents a number of physical as well as social limitations to what the user can do in conjunction with the meeting. Therefore, and given the recent, rapid rise of smartwatch devices, we envision that many users will soon want to bypass their phone and participate directly from wearable devices. As Pearson et al. suggest, there may also be value in using a wearable device like a smartwatch to deliver information to others in the immediate surroundings of a meeting as well [19].

Past work in glanceable and peripheral displays has shown this can be possible if the tool allows participants to interleave minor actions [17]. Researchers have put this principle to use to build interfaces for monitoring data or perusing documents while conducting another task [2,16]. Furthermore, work based on qualitative observation of people using multiple devices suggests the optimal way to integrate two devices may depend on whether the secondary device is being used in a parallel or sequential role [10]. Individuals often use a phone alongside other devices to perform support tasks or as a secondary, peripheral display presenting some other sort of information [14, 20]. However, as Hausen et al. [9] mention, digital devices often require our undivided attention, which can disrupt people working on a primary task.

To address this problem, prior work has looked at ways of using secondary devices to complement rather than distract from the use of a primary device. Chen et al. [3] explored ways of enabling interactions between two smart mobile devices, for example by using a smartwatch to check new emails or read text messages when the phone is not ready at hand. However, the majority of these projects have focused on a single user performing a solitary task.

In this paper we present two systems we built to address the challenges of divided attention and multiple tasks in distributed meetings for both collocated and remote participants (Figure 1). We also describe studies we conducted for both systems and use the results of those studies as the basis for implications for designing divided awareness interfaces for meetings.

NEEDFINDING STUDY OF DISTRIBUTED MEETINGS

In order to understand current practices with distributed, video-mediated meetings, we began with a series of surveys and exploratory interviews with 52 information workers across a variety of professions including technology (software developers and designers), engineering and banking. In these needfinding exercises, we asked participants to describe their remote collaborations and to explain the configuration, setup, and nature of recent distributed meetings they had been a part of.

Overall, we found that distributed meetings were relatively common, with 84% of respondents reporting having them at least once a week (and 40% having them every day). While using video conferencing tools to hold meetings and deliver presentations often worked “well enough,” several interviewees provided examples of cases in which the technology available did not make meeting at a distance fully seamless. Two main areas for improvement emerged: enhancing participation in distributed meetings through receiving and sharing information.

In work extending [18], we conducted a broad set of interviews with people using off-the-shelf tools for a variety of distributed work tasks, including design meetings and status updates, as well as distributed learning environments. In many cases, a group of people in one shared environment connected to one or more remote participants or groups. Collocated groups typically shared a single connection to the remote users (e.g., a webcam and large monitor on one end of a conference room with the participants clustered around the other end of the room).

Our observations of these groups revealed a set of both individual and collaborative behaviors that augmented the meeting. We saw users: research topics related to the meeting content on a web browser; pull up a document shared before the meeting to review and annotate it to later make a comment on it; find related content with the goal of pointing it out to a local participant; and capture content (usually with a mobile phone) created in the space on whiteboards or other surfaces to share with remote colleagues. All of this extraneous behavior was, however, done completely outside of the primary meeting software itself. Thus, there was no means to create general awareness of related information other people might be finding. Distributed members were particularly left out, unable to determine what people were discussing on the side in the remote space. Furthermore, none of these behaviors was in any way linked to any meeting recordings – all of this extra work done during the meeting itself was therefore lost and would often have to be recreated later.

Potential for mobile devices to enhance participation in distributed meetings

Use case 1: Helping remote people receive information on the go

Many videoconferencing tools now offer apps for smartphones and smartwatches to allow people to participate remotely. We examined reviews of one such tool (Cisco WebEx) in the iTunes and Google Play stores online to understand how people reported using the apps and what their experience of these were. Reviews left by users included mentions of calling into meetings via phone in a variety of on-the-go situations (when in a parking lot, traveling, commuting, joining on mobile while on the way to a meeting and until a person could get to the office). Other scenarios for using the apps included using the phone as a second or third screen to augment a laptop or desktop, or in cases where it was more convenient to bring a portable (phone) device as opposed to a laptop.

However, while reviews pointed to the convenience of calling in via the mobile phone, some users remarked that they were not able to participate as desired in the main meeting. For example, there was "*no way to participate in chat room*" (Zoom.us review, Google Play store comment from December 2015) and more importantly, "*nice way to attend meetings on the go but did not find it suitable for presenting*" (Zoom.us review, Google Play store comment from November 2015).

Use case 2: Helping remote people receive information about local artifacts and sketches

In our exploratory interviews, we found a need for remote participants to be aware of physical objects, such as paper designs or sketches that were present and visible to the group in the main room. This use case often came up for groups of designers, who often printed and hung up sketches and mockups around the room for others to view and comment on. However, when one designer was out of the room, this common practice broke down: *You could just [draw a sketch] in the notebooks and then I send pictures of the notebooks and I don't have to capture what was on the wall.*

Another example of needing to share information with remote participants came from a mobile app designer who would hold up a phone to the laptop's webcam, as this was the easiest way to share awareness of an interaction with a physical device (the phone).

I deploy the application on my phone. Then we use the videoconferencing so I can put the phone in front of my laptop camera so you can see what it looks like.

However, the problem with this method was that the information being shared was ephemeral and it was not possible to refer back to that example later if needed.

Use case 3: Helping collocated participants share information and contribute

Finally, interviewees also described situations where collocated participants had a need to be able to share and converse within their local context. A common configuration for remote meetings was one shared computer broadcasting the meeting to several people who might be gathered around it in the same room.

The Macedonian team, it really depends on how many people join. If all three of them join, then they usually use just one laptop and the three of them sit together.

This laptop, or common meeting broadcast system, might be a permanent fixture of a meeting room or in other cases it could be a personal device from someone in the group:

We'll pull up guys in a hangout. What we normally do. In order to do that you use the laptop that's on the table, and the laptop on the table is not always great but sometimes they bring their own laptop or plug that in and that's normally better.

Many interviewees highlighted how this configuration made it very difficult for other participants in the same room to participate in the meeting through screen sharing because they typically did not connect to the meeting directly themselves:

Say we were in video and Julia wanted to show it quickly, it's probably not her computer that we're actually connected to so and because it just wouldn't be a quick simple thing to do, we probably wouldn't do it. The person's computer, they would try to bring up the file, so it would be, in an ideal

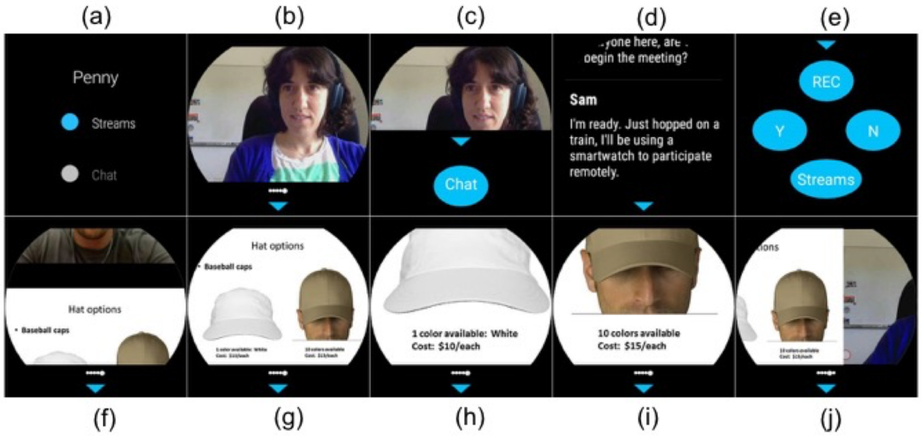


Figure 2. MixMeetWear views. *First row:* (a) MixMeetWear main menu; (b) Pressing "Streams" opens a view displaying a keyframe from the first stream; (c) Dragging the triangle at the bottom opens the control panel; (d) Touching the chat button opens a view displaying messages; (e)Touching the triangle at the bottom opens the chat control panel. *Second row:* (f) Swiping a keyframe moves to a different participant's stream; (g) In this stream the participant is shares a slide; (h) Double tapping zooms the view; (i) Zoomed-in keyframes can be panned; (j) Swiping horizontally moves to a different keyframe within a stream.

situation, you would just, Julia would do something and boom, her screen would be shared. She doesn't even connect to the meeting, usually, through her computer. It's usually just that one computer.

The inability for a person in a collocated group to share information into the main meeting often meant they had to use a workaround, causing delays and disruption while the material or file is transferred to the host's laptop.

Sometimes, we have issues where the presenter might not be able to present things. In that case, we just transfer the presentation to a laptop that the organizer [host] has and then he or she presents it from their laptops. That happens sometimes as well.

As one solution to this problem, using mobile devices connected through an Apple TV and Airplay was reported by one group. However, few of our interviewees reported having or using an Apple TV in their meeting setup.

but the Mac can't do that. It can't send a link to the Apple TV even though the iPad and the iPhone can. So in general, when it's time to show a video, we'll beam it off a phone or an iPad. We won't do it off a laptop.

Use case 4: Supporting divided attention

Finally, interviewees also explained situations in which their attention might be divided during a meeting, either for work-related tasks (e.g. searching related material, having a sidebar conversation with a coworker around something that was mentioned, or simply doing other unrelated work).

Sometimes, like Julia, for instance, almost always brings her laptop to the table and is usually multitasking in the middle of the conversation.

Right, so we have this chat window here, we have one for the whole meeting so you can put conversations or questions there if you wanted, but you can also have one on one conversation with other people. A lot of times, what happens, they'll bring up a sustaining question, this feature doesn't work and so the developers will have a back conversation. You all look at it, I will look at it, that kind of thing.

This implies a need to support individuals in either doing something related to the meeting or not, while enabling them to maintain a broader awareness of the conversation and/or allow them to catch up on what they might have missed later.

We addressed these various needs through the design of two mobile systems that link a user to an ongoing video conference: One through a smartwatch (MixMeetWear), which enables divided attention and participation on the go, and one through a mobile phone (MixMeetMate), which provides additional participation abilities to previously limited participants. We describe the design and evaluation of these two systems in turn.

THE MIXMEETWEAR APP

A smartwatch is less obtrusive and faster to access than a smartphone, making it more appropriate for peripheral meeting awareness [6]. While some commercial meeting systems have rudimentary smartwatch apps, none of them broach the difficult design challenges of interacting with and displaying persistent visual information (beyond, for example, a notification) [21]. MixMeetWear is an Android-based smartwatch application we designed to extend WebRTC-based meeting systems that support live audio, video, and chat. In particular, we built MixMeetWear to extend MixMeet, a web-based meeting system that also performs live content analysis and archives important keyframes from each participant's stream on-the-fly [1].

Our goal is for MixMeetWear to run independently on a smartwatch. However, current smartwatch hardware requires that networked applications send and receive data via an associated mobile device. Therefore, MixMeetWear currently consists of two applications, a mobile app and a wearable app. The mobile app interoperates with MixMeet to join meetings, handle the data between the server and the watch, and streams audio from the server. The mobile is not directly needed during a meeting. The wearable app displays visual content from the meeting and allows for simple input from the user.

The wearable app takes advantage of the keyframes generated by the MixMeet system. Keyframes are screen captures of a user's shared screen during meetings, which are stored for later reference. Keyframes are generated in the server, and parameters used are varied based on what a user is sharing. For instance, when sharing a standard webcam view, a certain level of change in a view is required for a new keyframe to be generated in order to avoid redundant uninformative frames displaying (e.g., a user's face). On the other hand, when sharing a slide show, for example, any change in view will generate a new keyframe.

MixMeetWear's UI includes a main menu and two other views; the stream view and the chat view (Figure 2). The main menu offers access to both views, however users can also move between the two views without having to go back to the main menu. Both views contain a control panel that can be opened by pressing or dragging an icon at the bottom of the screen.

The stream view consists of a series of keyframes that can be scrolled, zoomed, and panned. Keyframes are sent over to the watch whenever a new keyframe is generated in MixMeet for any meeting participant. On the watch, keyframes are divided into user-specific streams. Swiping up and down changes the selected user's stream, and swiping left and right cycles between keyframes within a particular user's stream. Double-tap zooms in and out of a keyframe. A zoomed-in frame can be panned freely. Using the control panel, users can open up the chat view.

MixMeetWear also includes essential mechanisms enabling the watch-based participant to interact with the meeting's separate chat channel. The app supports basic yes/no responses as well as open-ended speech-to-text-based comments. Both types of responses are triggered via the control panel. Yes/no responses are sent directly to the chat after pressing the corresponding button.

The speech-to-text function displays a view instructing the user to speak and then displaying the interpreted result. Afterwards, a confirmation timer is displayed, allowing user time to cancel the action. When the timer finishes, the result is sent to the chat.

MixMeetWear Evaluation

To better understand the use and usefulness of various smartwatch app functionalities, we conducted a study with 12 participants (7 male, 5 female). The participants' *meeting-based* task was to use our smartwatch application to participate in a meeting, while also carrying out two *parallel* tasks on a laptop (one task followed by another task). The 10-minute long meeting was pre-recorded with three members, and it was simulated on the watch as a live meeting. The meeting's topic was to discuss ideas for new items that could be sold in a company's virtual store. In order to encourage study participants to look at both slide and webcam keyframes, the meeting included content sharing alongside the audio conversation. Two different meeting members used screen sharing to show slide content relating to the discussion. The third person held up a physical water bottle to the webcam to illustrate a product design.

The participant's meeting-based task was to stay aware of the meeting, trying to pay attention to what was being discussed and what was being shown on the watch. Furthermore, at key points during the meeting, input was expected from participants. At these points, it was up to the participants to react to the situation, using the watch app's chat functionalities the way they saw fit.

The meeting was divided into two parts of roughly equal length, one for each parallel task. The *low attention* parallel task simulated traveling in a metro. A timed slideshow was running on a laptop, displaying different metro stops one at a time. The time a particular stop was being displayed varied between 30 and 60 seconds. The participant's task was to get off at the correct stop, which they signaled by hitting escape on the keyboard and closing the slideshow. Regardless of whether participants got off correctly or missed the stop, the task continued until that part of the meeting ended.

The *high attention* parallel task involved users preparing a short slide show that was based on a template. In the template, users were asked to present different types of bags and watches that they liked and describe them in the form of pictures and text. Again, the task ended when the corresponding part of the meeting was finished.

During the meeting, keyframes were sent to the watch at a varying pace. Webcam frames for a single stream were usually updated roughly once every 20 to 30 seconds, however at key points during a meeting, when e.g. there was a major change in webcam view, the frames would update once every 5 to 10 seconds. Slide frames always updated right away when there was a change in a shared slide show, i.e., when a meeting participant moved to the next slide.

Procedure

Prior to starting the meeting, participants were walked through all the functionalities of the watch so that participants could try out each to practice and familiarize themselves with the app. The order in which the parallel tasks were presented was counterbalanced, 6 participants starting with the low attention task and another 6 with the high attention task.

After each parallel task, participants filled out a paper version of the NASA-TLX scale and answered some comprehension questions about the contents of the

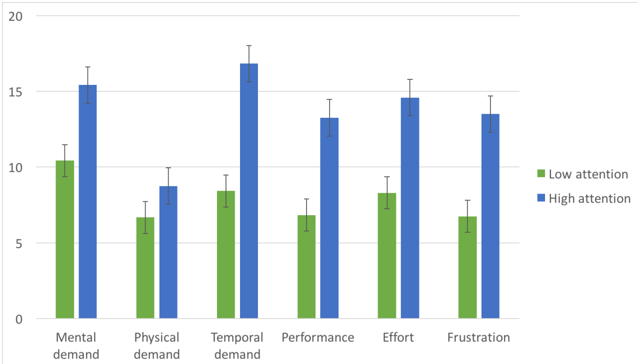


Figure 3. Mean ratings for TLX scale items in low and high attention tasks (Higher score = more demand/difficulty; max=21).

completed part of the meeting (such as who said what, and what appeared on the slides).

Participants answered additional questions after completing both tasks. Furthermore, to measure participants' subjective experience and attitudes towards the smartwatch app, we presented a set of Likert scale and open-ended questions focusing on general opinions about what they liked and disliked about using the app to follow the meeting.

Setup

We used the LG G Watch R smartwatch, paired with a Motorola Nexus 6. For the audio, a pair of high quality earbuds were attached to the phone. For the parallel tasks, participants used a Windows 7 laptop and a standard optical mouse.

Results

We gave participants a score for their performance in the meeting comprehension task as well as the high attention parallel task. Three researchers graded all participants independently and unaware of the grades given by others, and the mode of the scores was used in the analysis.

Performance on meeting comprehension task

We measured participants' awareness of the meeting using their answers to meeting comprehension questions. The maximum score was 19. Overall, participants performed well in the main task. The lowest score by a participant was 12, and the highest was 18. The mean score was 16.3 (SD = 1.72). There was no statistical difference between different task orders.

Performance on parallel tasks

All participants finished the low attention task correctly (that is, they paused the meeting when they arrived at the correct metro stop). As for the high attention task, we graded participants based on how many elements in the slides they had filled out. Elements counting towards completion were e.g. items in bulleted lists, cells in tables, and pictures. The theoretical maximum score was 20. The lowest score by a participant was 2 and the highest was 10. The average was 6.5 (SD = 2.78).

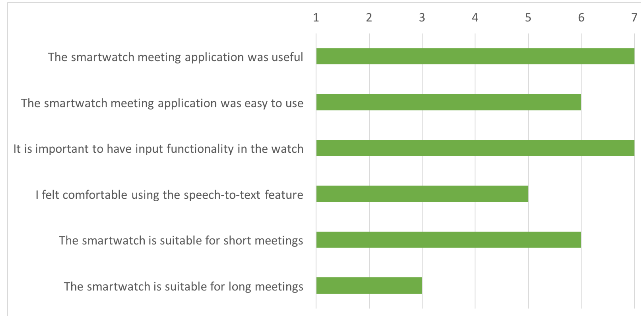


Figure 4. Median ratings from the UX questionnaire.

Subjective results

Means and standard deviations from the NASA TLX for both parallel tasks are presented in Figure 3.

A paired-samples t-test shows significant difference in all but physical demand. Participants found the high attention task to be more mentally demanding, t(11) = -6.124, p <.001, as well as more temporally demanding, t(11) = -4.994, p <.001. Participants also felt they performed more poorly, t(11) = -3.519, p = .005, had to put in more effort, t(11) = -5.100, p <.001, and were more frustrated, t(11) = -4.521, p <.001, during the high attention task.

General and open-ended questions

Participants also answered questions about their overall experience and attitudes in a 7-point Likert scale, where 1 was "strongly disagree", 4 was "neither agree not disagree", and 7 was "strongly agree". The results are presented in Figure 4.

The smartwatch was seen as clearly easy-to-use (median = 6) and useful (median = 7). Chat input functionalities were seen as important (median = 7). The speech-to-text function was less favored due in part to various social concerns (median = 5). One participant explained:

I would feel awkward using speech-to-text in public. Unlike a regular phone conversation, where I would probably be talking often, I would only occasionally say something while using the watch, which may be strange to those around me.

Another participant mentioned a perceived disconnect between the two different modalities of responding:

When someone is invited to participate by voice, I feel like I need to respond by voice. I think it would be more appropriate if the chat was under a question and answer format. The question the presenter is asking is written out and I answer y/n, instead of answering text to a voice question.

Additionally, the smartwatch was seen as suitable for short meetings (median = 6), but for long meetings opinions were less favorable (median = 3).

Lastly, we asked participants open-ended questions about what they particularly liked and disliked about using the smartwatch as a way to stay aware of a meeting, and in what kinds of situations they would feel it to be most suitable. The most common reported positive trait was that the watch app *allowed staying aware of the meeting* while doing something else (6/12). Other common answers included *being able to see the slide keyframes* (3/12) and *being able to respond* via the chat (3/12).

The most common complaint was that users still *needed to manually switch streams* when for example another person started presenting slides (4/12). Some participants specified that they wished the stream would automatically change based on which participant is speaking. Other complaints were that *responding required too many steps* (i.e., need to switch views, open the panel and press a button) (3/12), and that the watch resulted in *physical challenges*, i.e., difficult to type with the watch hand or need to lift arm up to see screen properly (3/12).

Interactions

We recorded all smartwatch interactions during the study. We were particularly interested in participants' behavior regarding zooming and panning, i.e., whether they were seen as useful functionalities and whether those would be used in conjunction with a parallel task.

A total of 49 key frames were shown on the watch by the end of the meeting, 6 of which were slide frames. Log data shows that only 3 keyframes (all of which were slides) were zoomed and panned by multiple users. Additionally, 2 webcam frames were zoomed in once by one user. Other key frames were not zoomed in at all.

We also looked into participants' response behavior. Speech-to-text usage was rather low as expected. On average, participants had one speech-to-text response, while three participants did not use speech-to-text at all. On the other hand, predefined yes/no responses were used much more often than we expected. Users had an average of 4.25 responses per meeting (SD = 2.01), the minimum amount of responses being 2 and the maximum 9.

Summary

The MixMeetWear app gave insight into feasibility and participation issues in taking part in distributed meetings via smartwatch. This addressed two use cases from our needfinding: Use case 1: Helping remote people receive information on the go, and Use case 4: Supporting divided attention. In order to address the remaining use cases, we designed a smartphone app to allow participants to contribute more easily and more richly to a hub-and-spoke meeting.

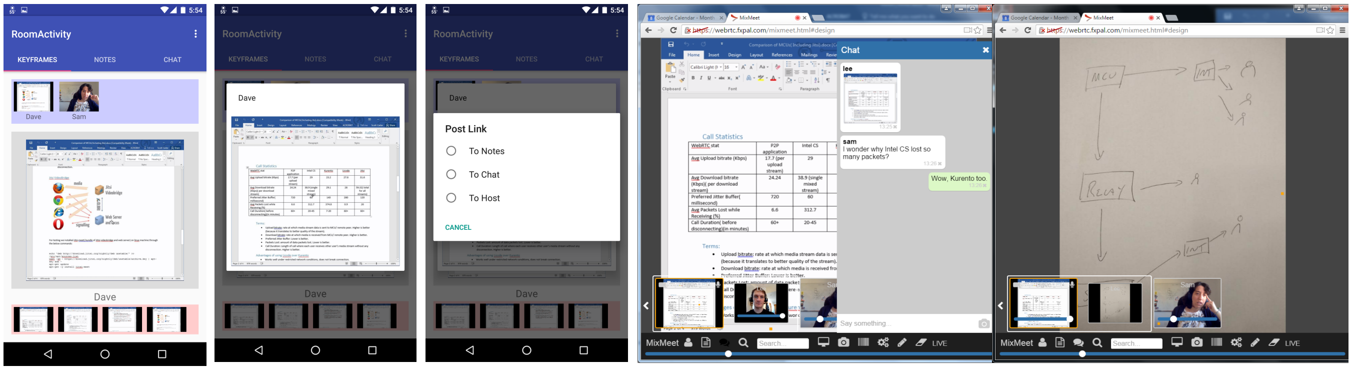
THE MIXMEETMATE APP

As noted earlier, our needfinding work revealed that hub-and-spoke arrangements in distributed meetings have multiple issues. First, participants in the collocated group who are not the host often want to refer back to previously presented material while not breaking the flow of the meeting, but have no means to do so. Furthermore, collocated participants will often share content with one another in side conversations, but these conversations are not available to the remote participant, nor is any related content archived for later reference or reuse. Finally, collocated participants will often use whiteboards, flipcharts, or other objects that are not visible to the remote participant. We observed some instances of collocated participants trying to take photos of this content to send to the remote participant in a backchannel, but since they did not want to interrupt the flow of the meeting these notes were sent using a separate application that again would not be archived with the rest of the meeting.

To address these issues, we developed MixMeetMate, an application designed to augment an ongoing teleconference. Since we also found that most meeting participants tend to have their phone with them anyway, we developed two versions of the app for mobile devices: one native Android and one web-based for iOS devices. Similar to MixMeetWear, MixMeetMate is designed to be used with the MixMeet WebRTC-based web teleconferencing system. With MixMeetMate, collocated users can quickly connect to a local host of a meeting-in-progress by either scanning a barcode on the host's MixMeet instance or by selecting a host from a list. MixMeetMate provides access to meeting content without initiating any new video or audio streams. In this way, MixMeetMate users avoid microphone feedback issues that occur when multiple users in the same space each connect directly to the meeting. It also minimizes local network load.

Profile photos of participants connected via MixMeetMate appear in the meeting's thumbnail participant list associated with the host to which they are connected. This allows participants to see at a glance which users are collocated and which are remote. MixMeetMate provides quick access to archived keyframes from each meeting participant's stream. They can save important keyframes to their notes or send them to the meeting chat. MixMeetMate also allows users to snap photos and upload them directly to their own stream or to the chat (Figure 5).

MixMeetMate users can also load links directly into the meeting. Importantly, all of this information is archived alongside the recording of the meeting itself, so that any users reviewing the meeting have access to all of its associated content.



(a) (b) (c) (d) (e)

Figure 5. MixMeetMate user Lee is viewing the document that Dave is sharing (a). He scrolls back in Dave's keyframe history and finds a keyframe shared earlier (b). He sends this to the chat (c) and Dave and Sam add comments using the web UI (d). Lee can also snap a photo and send it directly to the meeting view (e).

MixMeetMate Evaluation

In order to test the features of MixMeetMate and see how users engaged with captured meeting material and contributed content through their mobile devices, we conducted a study with 10 new users (7 female, 3 male). We used a similar setup to Study 1, in which participants took part in a meeting discussing products for sale.

Method

Participants came into the lab in pairs. The experimenter began by introducing the MixMeetMate system and showing participants how to use it. This was followed by a training session in which participants were each given a Nexus 5X phone to use for the study and practiced using the app (participants used only the Android native app for the study). Once they were familiar with the various features, the experimenter sent one participant into another room to simulate a "remote" participant, while one participant stayed in the main room with the experimenter (to simulate a collocated participant). The experimenter connected both phones to the meeting, so each participant was able to view and participate in the meeting.

The experimenter then began by sharing slides explaining the participants' task, which was to spend 10 minutes answering a series of questions about a jacket to sell in the virtual store. This included URLs and text information about the jacket (such as price). To take advantage of the mobility of the phone user and their ability to move around easily and capture parts of their environment, the task also asked participants to send a picture of a physical example of the jacket (located in a different part of each of their rooms). The experimenter collected the information that the participants sent and compiled it in a grid that was being shared via screen-sharing. Since the grid was always visible, participants had the option of using keyframes to review the questions if they forgot what they were supposed to look for.

After the 10 minutes had ended, participants filled out a questionnaire containing the NASA-TLX to assess difficulty and effort. Similar to the MixMeetWear evaluation, they also used a 7-point Likert scale to rate the usefulness both of the app as a whole and of individual features (such as the ability to browse keyframes, share photos, and share links). They then answered open-ended questions concerning what they liked and disliked about the mobile meeting system.

Results

Figure 6 shows the overall NASA-TLX ratings for performing the task. Compared to the ratings in the previous MixMeetWear evaluation (Figure 3), there were no significant statistical differences between the six ratings given by participants in the MixMeetMate task and the MixMeetWear-low attention task. This suggests that these two means of participating in a meeting (via smartwatch or phone) were viewed as equivalent in terms of their demand, performance, effort and frustration.

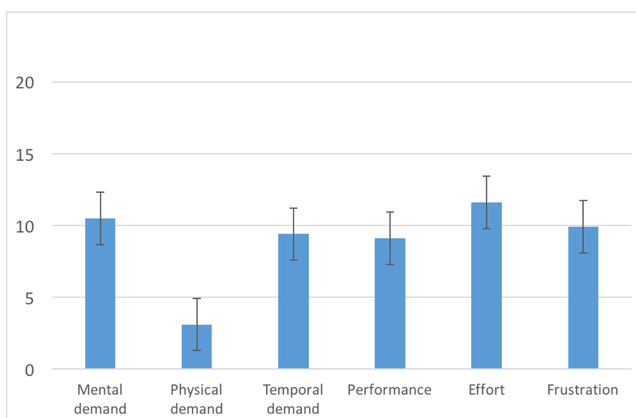


Figure 6. NASA-TLX ratings for mobile meeting task (Higher score = more demand/difficulty; max=21)

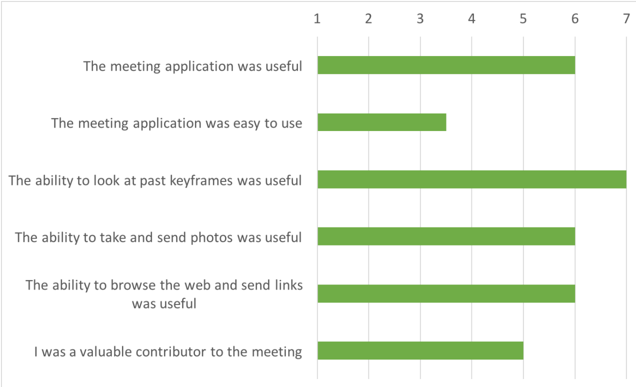


Figure 7. Median values for the UX questionnaire (1=completely disagree, 7=completely agree)

Figure 7 presents the median ratings of agreement with each of the follow-up questions regarding features and overall opinions of the application.

General and open-ended comments

We analyzed the general and open-ended comments for themes pertaining to three areas: (1) Perceived usefulness of the app in general, (2) Positive features and (3) Areas for improvement of the design.

The concept of the mobile app was, overall, seen as useful (median rating 6 out of 7). Many participants saw the utility in the standalone mobile phone app (the "remote" use case) and mentioned scenarios in which it would be useful. For example, two participants mentioned that it would be useful in cases where they were running late to a meeting and would need to join and participate on the mobile phone until they arrived in a more stable setting. Another participant appreciated having the integrated functionalities all in one place, mentioning that in her current group meeting setup, she had to have multiple different tools open to achieve the goals of information sharing with the group (including Slack, Google Drive, and Google Hangouts).

The ability to browse and share past keyframes was seen as one of the most useful features, as indicated by its high median score of 6.5. Participants expanded upon the utility in their open-ended comments. Four participants remarked on the benefits of being able to refer back to past material: "*Keyframes [were] very good for looking back at the task list, and keeping an eye on how the information I contributed was being synthesized*." Another participant mentioned that the ability to look back at past slides meant that he did not need to ask any questions of the experimenter regarding the task or take notes on the information requested: "*I didn't ask any questions because I knew I could go back*." At the same time, two participants also remarked that since the past keyframes were not zoomable, it was somewhat hard to read them due to the small size of the text on the slides.

Another positive feature of MixMeetMate was the ability to switch easily between the meeting interface and the web browser. In order to search the web from other mobile video-conferencing tools, a user needs to toggle between different applications (the video conference and the phone's web browser). In contrast, MixMeetMate allowed people to quickly and easily share links into the meeting without the extra steps of opening a new app, searching, copying and pasting the link back in the meeting itself.

Finally, the chat was viewed as a "*useful semi-permanent place to store all information*." Since MixMeetMate also supports choosing where to send information (to chat vs. to host), we asked participants to explain their decision for where they sent pictures. A majority of participants (7/10) sent pictures into the chat rather than to the host. One participant explained doing this because they felt the group was collecting all the information in the chat, and she could check what she had already contributed by scrolling back through the past chat history. She also mentioned that if the group had been discussing an item in further depth, she would have sent it to the host so that all participants could make it their focus. This suggests that having multiple places to send a piece of information allowed the senders to make nuanced decisions about the purpose and retrievability of shared information.

The meeting application, overall, was not rated as highly easy to use (median=3.5), although participants were able to grasp the concept of the different functionalities after the training. The main area of diminished usability reported by participants was the web browsing functionality (particularly issues with being able to go back in the web browsing history). Suggestions for improvement focused around overall streamlining the browsing experience (making the transition from web browser to meeting and back even smoother). Additional comments regarding changes to functionality included enabling one-on-one chat in addition to group chat (2/10), and making shared content more legible (2/10). We also noted that many participants sent a link in the chat and then followed up with text explaining the link (see Figure 8). One participant suggested that it would be useful to show previews of links to reduce the need for such manual annotation or description of the link content.

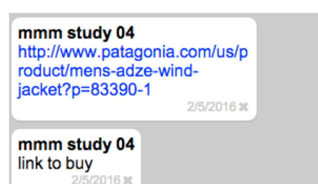


Figure 8. Example of URL sent to chat (top) and further explained by user in chat message (below)

DISCUSSION OF MOBILE MEETING PARTICIPATION

Participants were successfully able to use MixMeetMate to participate in a video meeting, either as a remote attendee or an associated collocated member. In particular, users appreciated the fact that the interface combined multiple different functionalities into one interface (current systems require multiple steps to share a URL while in a teleconference, including exiting the teleconference app in order to search the web, copying the URL, returning to the teleconference app, and pasting the URL in a chat feature). In the MixMeetMate app, completing such an action is more straightforward.

In both MixMeetWear and MixMeetMate, participants also positively reacted to the ability to go back to archived keyframes of a live meeting. These are features that are currently not available in commercial teleconferencing apps, such as Skype or Google Hangouts, and as we saw, were appreciated particularly for the use case of following a meeting while traveling or running late. Furthermore, many of the usability issues we found are either a result of a truncated learning curve (for example, one-to-one chat is actually already implemented but we did not cover it in the training session for timing reasons) or easily addressed (adding more support for integrating browser-based content).

The responses to the MixMeetWear app were more nuanced since we were investigating not only how well a smartwatch streaming application would allow users to stay aware of a meeting while focusing on additional simultaneous tasks, but also how users experienced the watch and how they interacted with it.

Based on users' performance in both the meeting and parallel tasks as well as their feedback, we conclude that the smartwatch can be a useful tool for staying aware of meetings. The watch app was generally seen to be both useful and easy to use. More optimal situations for using a watch are those that require low levels of mental effort, such as traveling or physically moving. Otherwise users' performance in the parallel task is significantly lowered. This is also supported by the fact that participants felt the watch to be more useful in short rather than long meetings.

Input functionalities were also seen to be very important, however user's feedback suggests that responses from watch users should not be expected often, but rather the focus should be on passive awareness. Further, using speech-to-text divided opinions between participants, and did not always work correctly. On the other hand, participants used predefined yes/no responses considerably more often than we anticipated, as many participants tried to express their opinion on matters being discussed in the meeting, even when input from them was not explicitly expected. The results might suggest that instead of using speech-to-text, a larger set of predefined responses could be introduced, or live audio should be streamed directly for others to hear.

Implications for design

Keyframes

The keyframe browsing was appreciated by both the watch and mobile phone participants. In the case of the watch, slide type keyframes were regarded as considerably more useful than their webcam counterparts. If utilizing a similar keyframe approach, we recommend live content analysis in the generation of keyframes. Including too many non-relevant webcam keyframes in streams makes it easy to accidentally overlook other, more relevant content and forces users to pay more attention to the watch, which should primarily act as a peripheral device.

To further streamline the watch app and emphasize the watch's role as a peripheral device, automatic switching between streams should be considered. This could be achieved by analyzing which user is currently the most active (e.g., speaking), and displaying that participant's stream.

Participants did less zooming and panning than we expected. We were correct, however, in assuming that slide frames would be the focus of such interactions. It is notable that the three frequently zoomed in frames were the only frames that contained a) pictures (more than one in every frame) and b) smaller text than other slide frames. This indicates that watch users prefer to glance at the frames without explicit interactions, but will use tools available if they need to take a closer look. Even without extensive use, being able to zoom in is an important feature, as we argue it is not realistic to assume that people would alter the way they produce slide shows and other similar content just for the sake of small-screen users, by e.g. using larger font-sizes than normal.

While our study used the smartwatch form factor as an example of a lightweight, unobtrusive device, future work can investigate the role of other types of wearable technologies for peripheral meeting awareness. For example, a head-mounted display could provide another way of presenting glanceable information about meeting activity, although the means of interacting with the content may involve different techniques. Another promising area for future research is to explore additional cues (such as haptic alerts) that can further help passive meeting participants know when to direct their attention to the device (for example, when something is likely to require their input, their name is mentioned, or a change in slides is detected).

Choices about how and where to share information

Beyond improvements to the usability and UX of the MixMeetMate app, participant reactions to the tool suggest that the integration of slide review, chat, and screensharing created a seamless repository for meeting-related information that was easy to navigate. At the same time, participants expressed a feeling that the "chat" section might not be the best place to store information from multiple contributors: "*Not sure that chat is the best mechanism if lots of people contributing – info seems easily lost?*" The challenge that comes along with enabling mobile users to contribute information to the meeting is an increased volume of content that needs to be stored, arranged, and made sense of. The degree to which multiple participants contribute items may depend on the nature of the task and size of the group; nonetheless, finding additional ways to prevent overload of information (particularly to viewers on a small screen) is of interest for future work.

CONCLUSION

The "anywhere, anytime" meeting trend means that more people in the future will participate in meetings from remote devices, often while engaging in a variety of other activities. In many cases, these activities are work related and have direct links to the meeting at-hand. However, off-the-shelf tools often make it difficult to integrate this "shadow work", leaving it siloed on participant's devices. In this paper, we began with insights gleaned through needfinding that identified ways in which wearable and mobile devices can help meeting participants contribute and send information when on the go or when not connected to a host's laptop. Initial deployments with a meeting-based app in two contexts (on a smartwatch and on a mobile phone) suggest that users may value and utilize customized features such as the ability to review past material, as well as to decide how and where to send it. Insights from this work can be applied to future tools designed to help bring wearables and mobiles into meetings even further.

REFERENCES

1. Scott Carter, Laurent Denoue, and Matthew Cooper. 2015. Searching and browsing live, web-based meetings. *Proceedings of Multimedia*, ACM, 791–792.
2. Scott Carter and Laurent Denoue. 2009. SeeReader: An (Almost) Eyes-Free Mobile Rich Document Viewer. *arXiv preprint arXiv:0909.2185*.
3. Xiang'Anthony' Chen, Tovi Grossman, Daniel J. Wigdor, and George Fitzmaurice. 2014. Duet: exploring joint interactions on a smart phone and a smart watch. *Proceedings of CHI*, ACM, 159–168.
4. Morten Esbensen, Paolo Tell, and Jakob E. Bardram. 2014. SideBar: Videoconferencing system supporting social engagement. *Proceedings of CollaborateCom,* IEEE, 358–367.
5. Four Trends Driving Cloud Video Conferencing in 2014. https://community.lifesize.com/docs/DOC-1190.
6. Wayne CW Giang, Liberty Hoekstra-Atwood, and Birsen Donmez. 2014. Driver Engagement in Notifications A Comparison of Visual-Manual Interaction between Smartwatches and Smartphones. *Proceedings of HFES*, Sage Publications, 2161–2165.
7. Andrew Greaves and Enrico Rukzio. 2009. View & Share: Supporting Co-present Viewing and Sharing of Media Using Personal Projection. *Proceedings of MobileHCI*, ACM, 44:1–44:4.
8. Drew Harry, Eric Gordon, and Chris Schmandt. 2012. Setting the stage for interaction: a tablet application to augment group discussion in a seminar class. *Proceedings of CSCW*, ACM, 1071–1080.
9. Doris Hausen, Aurélien Tabard, Attila von Thermann, Kerstin Holzner, and Andreas Butz. 2014. Evaluating peripheral interaction. *Proceedings of TEI*, ACM, 21–28.
10. Tero Jokela, Jarno Ojala, and Thomas Olsson. 2015. A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks. *Proceedings of CHI*, ACM, 3903–3912.
11. Tero Jokela and Andrés Lucero. 2014. FlexiGroups: Binding Mobile Devices for Collaborative Interactions in Medium-sized Groups with Device Touch. *Proceedings of MobileHCI*, ACM, 369–378.
12. Andrés Lucero, Jussi Holopainen, and Tero Jokela. 2011. Pass-them-around: collaborative use of mobile phones for photo sharing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1787–1796.
13. Andrés Lucero, Jaakko Keränen, and Hannu Korhonen. 2010. Collaborative use of mobile phones for brainstorming. *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, ACM, 337–340.
14. Sang-won Leigh, Philipp Schoessler, Felix Heibeck, Pattie Maes, and Hiroshi Ishii. 2015. THAW: tangible interaction with see-through augmentation for smartphones on computer screens. *Proceedings of TEI*, ACM, 89–96.
15. Hajin Lim, Hyunjin Ahn, Junwoo Kang, Bongwon Suh, and Joonhwan Lee. 2014. Ubi-jector: an information-sharing workspace in casual places using mobile devices. *Proceedings of MobileHCI*, ACM, 379–388.
16. Tara Matthews, Anind K. Dey, Jennifer Mankoff, Scott Carter, and Tye Rattenbury. 2004. A toolkit for managing user attention in peripheral displays. *Proceedings of UIST*, ACM, 247–256.
17. Tara Matthews, Tye Rattenbury, and Scott Carter. 2007. Defining, designing, and evaluating peripheral displays: An analysis using activity theory. *Human–Computer Interaction* 22, 1-2, 221–261.
18. Jennifer Marlow, Scott Carter, Nathaniel Good, Jung-Wei Chen. 2016. Beyond Talking Heads: Multimedia Artifact Creation, Use, and Sharing in Distributed Meetings. *Proceedings of CSCW,* 1703-1715.
19. Jennifer Pearson, Simon Robinson, and Matt Jones. 2015. It's about time: Smartwatches as public displays. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 1257–1266.
20. Stephanie Santosa and Daniel Wigdor. 2013. A Field Study of Multi-device Workflows in Distributed Workspaces. *Proceedings of Ubicomp*, ACM, 63–72.
21. Cheng Xu and Kent Lyons. 2015. Shimmering Smartwatches: Exploring the Smartwatch Design Space. *Proceedings of TEI*, ACM, 69–76.