

# Experiences2Go: Sharing Kids' Activities Outside the Home with Remote Family Members

Kori Inkpen<sup>1</sup>, Brett Taylor<sup>1,2</sup>, Sasa Junuzovic<sup>1</sup>, John Tang<sup>1</sup>, Gina Venolia<sup>1</sup>

<sup>1</sup>Microsoft Research

One Microsoft Way

Redmond, WA 98052 USA

{kori | sasajun | johntang | ginav}@microsoft.com

<sup>2</sup>University of Saskatchewan

110 Science Place

Saskatoon, SK S7N 5C9 CANADA

brett.taylor@usask.ca

## ABSTRACT

Video communication is moving beyond face-to-face discussions on desktop computers to sharing experiences out in the real world. We explored how mobile video could enable distributed family members to share experiences wherever they occurred – kids' sporting events, birthday parties, etc. We investigated how people used two technology probes to share activities outside the home: an iPad running Skype and our Experiences2Go prototype composed of a networked slate and a camcorder on a tripod. We observed their use in the field with nine families and explored the impact that their mobility, optical zoom, and multiple view features had on sharing the experience. We identified four sets of stakeholders in sharing experiences, the variety of sharing scenarios enabled, and reactions to the features that each probe offered, leading to design considerations for future mobile shared experience systems.

## Author Keywords

Shared experiences; family; technology probe; CSCW.

## ACM Classification Keywords

H.4.3 [Information Systems Applications]: Communications Applications – computer conferencing, teleconferencing, and videoconferencing.

## INTRODUCTION

*Grandma feels like she is missing out on her grandkids' soccer games, swim meets, and play activities, since they live 200 miles away. Sometimes her daughter will call her afterwards to tell her how it went, but Grandma still feels a little left out since she is not able to be there. They tried using Skype on her daughter's phone, but it did not provide an adequate view of the activity and it was tiring for her daughter to hold the phone up the whole time.*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CSCW'13, February 23–27, 2013, San Antonio, Texas, USA.  
Copyright 2013 ACM 978-1-4503-1331-5/13/02...\$15.00.

Video chat is moving beyond just face-to-face discussion as people discover new ways to use video to socially connect with remote family members and close friends [2]. The previous narrative illustrates an increasingly popular type of social connection where family members remotely join an event that they cannot attend in person. Such situations occur frequently due to physical distance, availability (e.g., busy schedule), or ability (e.g., wheelchair accessibility). Many of these events occur outside the home, prompting people to use mobile video chat applications, such as Skype and FaceTime, to include remote family members.

Unfortunately, today's mobile technologies limit what types of activities people can easily share. Current mobile video chat tools are limited by the cameras on the devices, which typically do not offer sufficient zoom or depth of field to clearly capture events at a distance, such as a soccer game. Although mobile devices offer front- and rear-facing cameras, video chat applications require the local person to choose between the two cameras. As a result, the remote person can either see the event or chat face-to-face with the local person, making it hard to truly share the experience.

Our research explored how to enable sharing experiences over distance and identified what capabilities are important to provide that sense of sharing. We studied the use of two technology probes [4] in real-world situations to better understand problems with existing technologies and design opportunities for the future. We focused on kids' events that occur outside the home (e.g., park, school, gymnasium).

We studied the use of two systems for sharing experiences: 1) an iPad running Skype and 2) an Experiences2Go (E2G) prototype we built, consisting of a networked slate and



Figure 1. Mom remotely watches her daughter play together with Dad through the Experiences2Go prototype.

camcorder mounted on a tripod. The distinct capabilities of the two probes allowed us to test a variety of features. The iPad is highly mobile but can only transmit the front- or rear-facing camera view at a time. E2G can capture events at a distance using the camcorder's telephoto lens and provides the remote person with simultaneous views of the event and the local people. E2G's tripod mount allows it to stand without being held, but makes it less mobile.

We deployed the technology probes with nine families to understand their use in real-world settings, what types of activities they enable, and the current obstacles to providing engaging experiences. Our observations provided insights into the types of engagement made possible when people at a distance can share experiences together. Our experiences also showed that this scenario is more than streaming a live event – it is about socially connecting people.

This paper first presents related work. Then it describes a pilot study we conducted to make some initial design choices. It then presents the E2G prototype and results from our deployment with nine families. Finally, it provides insights into sharing experiences on the go and design considerations for technologies that enable them.

#### **RELATED WORK**

Traditional videoconferencing systems work well for face-to-face conversations but are limited in their ability to go beyond the talking heads experience. Specifically, they are not well suited for supporting many of the activities that naturally arise during a social interaction. Bly et al. highlighted the importance of both informal and task-specific interactions for organic, spontaneous collaborations [1]. Furthermore, Gaver et al. [3] found that talking heads did not provide adequate information about the remote environment, which is important when sharing an activity.

Despite the limitations of existing systems, people have used video in their homes for personal connections. Kirk et al. [7] surveyed home users and identified that a desire for social closeness motivated using video communication. Judge et al. [5] described how distributed family members used video to share experiences across distance. Brubaker et al. [2] found that people are exploring and experimenting with using video to share a diverse range of activities (e.g., watching movies together, cooking with mom, remote baby-sitting). People have also been using video outside the home to maintain personal connections. Juhlin et al. [6] analyzed how people use mobile technology to broadcast live video, enabling sharing experiences even when they are not at home. O'Hara et al. [10] found that over half of video calls made on mobile phones occurred outside the home or work settings, indicating the growing opportunities for video chat on the go. Peek-A-Boo [9] explored creating video connections from the home to a mobile phone.

Prior research has also investigated prototypes designed specifically for sharing experiences. Weisz et al. [15] created a system for sharing text messages while watching video together. Romero et al. [12] created a system for sharing

awareness of daily activities through text and pictures among remote family members. Vutborg et al. [14] created a system designed for remote communication between grandparents and their grandchildren. They found that augmenting the video chat with other shared context was important. This finding was supported by Raffle et al. [11] who designed StoryVisit, which combines video communication with reading a book together. Others have also created video chat systems that provide a shared communication context. Yarosh et al. [16] created a system for children to have play dates over a distance using video communication and a shared task space, and Mueller et al. [8] enabled a shared jogging experience between remote joggers just through a live audio connection and spatialized sound to create a sense of jogging side-by-side.

These studies and prototypes demonstrate the desire to push the boundaries of existing technology for personal video communication. People want to share experiences using video communication, both at home and in mobile environments. Our work differs from prior work in two ways. First, we investigate the notion of using video to share experiences outside the home. Second, rather than focusing on a single system for sharing experiences, we studied two technology probes with fundamentally different features to get a wider range of insights into pain points and solutions for them when sharing experiences on the go.

#### **PILOT STUDY**

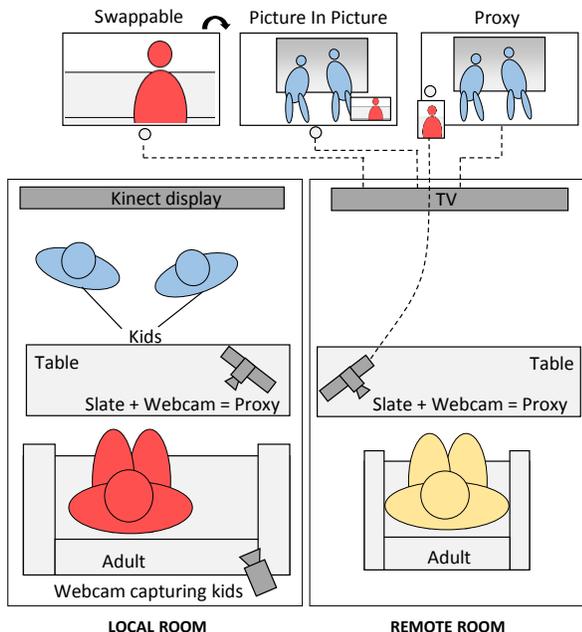
We began by exploring whether it is beneficial for remote users to be able to see the activity itself and interact with the local adult attending the event. However, it was not clear whether it was necessary to show views of the event and the local adult simultaneously, and if so, how to present these views to the remote attendee. We conducted a pilot study to help us understand the importance of multiple views and identify a good way to present them.

#### **Methods & Setup**

We recruited four groups consisting of two adults and multiple children. We looked for couples with kids or adults whose kids were friends with each other. Two groups were mother/father and two were mothers who were friends. We asked these groups to visit our lab and watch their children play Microsoft Xbox Kinect games.

The children played from a selection of Kinect games. One adult ("local adult") was seated on a sofa watching from behind as the kids played. A Microsoft LifeCam Studio webcam mounted behind the local adult captured a view of the children playing, similar to what the local adult saw. The second adult ("remote adult") was in an adjoining room watching the kids play via Skype on a 42" wall-mounted TV. Sound was transmitted using ClearOne USB speakerphones. The local adult always saw the remote adult on a Samsung 11" slate positioned on a table in front of the sofa, offset approximately 45 degrees to the right. This slate displays video captured by a webcam in the remote room. To capture video of the local adult, a second local webcam was mounted near the slate. We explored three arrangements for

presenting the local video feeds to the remote adult, which are depicted in Figure 2.



**Figure 2. Pilot study setup: (left) local and (right) remote rooms. In remote room, callouts show the swappable, picture in picture, and proxy conditions.**

In the *Swappable Views* condition (Figure 2), the remote adult could see on the TV either a view of the children playing Xbox or the local adult, but not both at the same time. Both adults could “swap” the views using a keyboard key. The local adult had an icon on their slate that indicated which view the remote adult was viewing. The remote adult was captured by a webcam centered under the TV.

In the *Picture-in-Picture* condition (Figure 2), the remote adult could see both the children and the local adult at the same time on the TV. A keyboard key was used to switch between having the children full-screen and the local adult in a small inset window, or vice versa. Similar to the first condition, the local adult knew which video view was enlarged by an icon that appeared on their slate. As before, a webcam centered under the TV captured the remote adult.

In the *Social Proxy* condition (Figure 2), we explored the use of a separate device to represent the local adult. The remote adult was provided with a separate screen (an 11” slate) to view the local adult, rather than having them appear on the TV. The slate was placed on the table directly in front of the participant, offset by approximately 45 degrees to the left. The webcam capturing the remote adult was placed directly above this slate to provide more accurate gaze information.

We ran each condition for approximately 10 minutes and concluded the pilot study with a brief interview.

### Results on Multiple Views

*Swappable Views.* People generally did not like swappable views. They did not like having to switch back and forth between the views and often lost track of what was

happening in the other view. There was also some confusion as to who should swap the views. The local adult knew what was going on in the activity, but did not know exactly what the remote adult wanted to see.

*Picture-in-Picture.* People generally preferred picture-in-picture. It enabled them to see both views simultaneously without forcing the remote adult to attend to two separate screens. Views were switched (event large vs. local person large) less frequently than in the swappable views case.

*Social Proxy.* While some people preferred the social proxy, others found it distracting to attend to two screens. They did not like having to turn their heads back and forth. However, several local people liked that they could easily tell which screen their partner was attending to given the gaze feedback provided in this configuration.

During this pilot, we observed that socializing between the local and remote adults was an important part of the experience. The adults frequently talked with each other while watching the children. As a result, we developed a prototype design that provided simultaneous views of the local adult and the activity.

### EXPERIENCES2GO (E2G) PROTOTYPE

Building on our pilot study results, we designed a prototype to explore mobile shared experiences with the following key design goals: 1) incorporate multiple simultaneous camera views; 2) use cameras that had pan, tilt, and zoom control and handled variable lighting conditions; and 3) work with standard Skype to leverage software with which our remote participants were already familiar.

Our E2G prototype utilized a Samsung 11” slate running Windows 7. We chose a slate, instead of the mobile phones used in earlier work, both to provide a comfortable size for viewing people over longer durations of interaction and provide a sufficiently powerful device to run our custom software described below. The front-facing camera on the slate provided the view of the local person. The built-in microphone and speakers on the slate were used for audio when possible; however, in many situations the ambient noise level was high, so the participants had to use a headset with a microphone.

Webcams are typically designed for near-field capture, which makes it difficult to capture events at a distance, especially since most webcams do not support optical zooming. Instead, we used a camcorder because its design affords capturing a wide range of activities. We utilized a Sony HDR-XR260V with a 30x optical zoom. Both the camera and the slate were mounted on a tripod to keep them steady without needing to be held (see Figure 3).

We streamed video from the camcorder to the slate using an EasyCap USB dongle, which converts a composite video signal to USB. We built a custom virtual webcam driver to produce a composite picture-in-picture image using the camcorder as the primary image and the front-facing camera on the slate as the secondary, smaller image.



**Figure 3.** The E2G prototype enables a remote person to attend an event while also socializing with others at the event.

Skype was used for the audio and video connection. A Verizon 4G hotspot provided internet connectivity. Thus, our E2G prototype was completely wireless with approximately 100 minutes of battery life, allowing portable use almost anywhere.

### FIELD STUDY

We deployed the two technology probes during real-life activities to explore their potential to share experiences with remote family or friends.

#### Methods

We recruited nine groups who were comfortable using Skype for video chat. Similar to our pilot, we recruited families with children who had a remote relative or close friend that wanted to watch an upcoming kids' activity. We provided suggestions but left it up to the parents to identify the event as well as the remote participant.

We wanted the experience to reflect real-life as much as possible, including the process of incorporating technology into an activity. Thus, we simply demonstrated how to use the devices, control the cameras, and made sure everything was working; otherwise, the participants did what they wanted and moved around with the devices as they liked.

After completing the study, we asked the local participant to complete a brief questionnaire and conducted a short semi-structured interview. The remote participant completed a similar questionnaire by email. We reviewed our interview and open-ended questionnaire response data for recurring themes and inductively compared specific instances of those themes across participants [13]. Our analysis focused on what features in the probes affected the participants' ability to share the experience over distance.

### Results

We first describe the activities we observed. Because of the diverse activities and configuration of people involved, the insights gained varied across the sessions. We briefly summarize, illustrating with a photo, each of the nine sessions. We then follow up with more general results aggregated across the sessions. Seven of the groups used both devices, but due to technical issues, Group 1 used only the E2G prototype and Group 2 only the iPad.

### Activities

The types of activities our participants chose for the study included watching children's sports events (swimming, gymnastics, soccer, and track and field), a birthday party, and play sessions (at the park and a local play center). Of the nine sessions, one involved remote grandparents, one a close friend, and the rest involved the other parent.

The activities can be clustered in terms of whether the adults were primarily spectators (i.e., just watching the activity), or if they were involved in the activity. Having the adults involved in the activity meant that 1) the local adult was interacting with both the children and the remote adult and 2) the children could also interact with the remote adult. We had three activities where the local parents were involved in the activities (a birthday party and two play sessions). The remaining activities were primarily spectator activities (sports) where the local and remote adults watched the activity together and talked with each other.

#### Group 1 - Fishing (local parent involved in the activity)

Our first session was a birthday party at a trout farm hatchery with two families and their five children. They connected with two grandparents who were located in different states. This group only used the E2G prototype. Although there was quite a lot of ambient noise from the pond (water pumps and a fountain), they did not use a headset since multiple people wanted to interact with the remote grandparents. However, using the internal speakers on the slate made it difficult for the local adult and kids to hear the grandparents unless they were very close to it.



**Figure 4.** Fishing with remote Grandmother and Grandfather.

The father was the primary person interacting with the E2G prototype and his parents (one at a time) were the remote participants. He spent a lot of time talking with his mother or father while watching the children. He gave his parents a tour of the facility and commentary about what the kids were doing. When one child caught a fish, remote Grandma was able to see this for herself from across the pond and became visibly excited: clapping and congratulating her grandson. From time to time, the grandkids would approach the E2G prototype and talk to their remote grandparents, even when neither of their parents were near it.

Both the father and the grandparents enjoyed interacting with the E2G prototype. They both also indicated that seeing each other was very important, as well as seeing the people and the activity at the same time.

When asked what they liked best about the experience, the father commented that “*having the [front-facing] and [rear-*

## Computer Supported Young People

facing] video simultaneously is a great advantage.” The grandfather explained that he liked “the opportunity to talk one on one with your partner, while feeling a part of the activities.” When asked what types of events the grandmother could imagine using a device like E2G, she replied, “Watching my grandkids. At games, vacations. Feeling like I’m there with them.”

### Group 2 – Park (local parent involved in the activity)

The Group 2 session involved free play in a public park on a sunny day. The mother attended with her two children (ages 18 months and 3 years) while the father worked from home. Because of the loud noise of lawnmowers in the park, the mother had to use headphones. Due to technical problems she was unable to try the E2G prototype.



Figure 5. Playing in the park with remote Dad.

The mother had a difficult time managing the iPad while attending to her children. She struggled to carry one child while keeping up with their toddler who was running around the park with a friend. She tried balancing the iPad on her stroller to free her hands, but this did not capture the view she wanted (and the iPad eventually fell to the ground). Her attention was fragmented between what was going on with her kids in the local environment and her conversation with her husband.

The mother commented that she wanted ultra-mobility with “a smaller device [she] could pull out of [her] pocket.” For her, the iPad was bulky. The father commented that he liked the ability to “be part of my kids’ life while I am not able to be there physically” even though he found the “shaky camera image” to be problematic. Neither parent felt that it was important to see each other.

### Group 3 – Play Center (local parent involved in the activity)

For Group 3, the father took his daughter (21 months old) to an indoor play center while the mother attended from home. The play center had music playing in the background, so the father wore a headset to hear the mother. The headset tethered him to the device, limiting his movement. He first used the iPad and later used the E2G prototype.

The father started with a tour of the space after which he followed their daughter around as much as possible using the iPad. Apart from the initial greeting, he primarily used the rear-facing camera. There was no interaction between the remote mother and the daughter.

February 23–27, 2013, San Antonio, Texas, USA



Figure 6. Play center with remote Mom.

After switching to the E2G prototype, the father struggled to keep up with his daughter, frequently moving and lifting the tripod. Eventually, he tried to position the tripod and engage in play with his daughter. He first positioned the camera so that the mother could see him push their daughter on a toy train, and later he aimed the camera to capture himself and his daughter playing with blocks on the floor. In both setups, he talked to the remote mother but did not seem to look at her on either the iPad or the E2G’s slate. He commented that the “[video of my wife] was not really that much important because I know her, I know where she is”.

The father felt that mobility was the most important feature, and therefore preferred the iPad setup, while the remote mother valued a zoomed in view of her daughter the most, and therefore preferred E2G. The mother also commented that she often had to ask her husband to adjust the camera to better capture their daughter and would therefore have liked “to be in control of the zoom in and zoom out feature” herself. The husband also agreed that it would be “an amazing idea... I’m busy with her [daughter] playing around... and of course she [wife] would be interested in seeing more.” Even though he was busy with the daughter and equipment, he said that “I didn’t really find it a big problem of like really concentrating on the conversation with my wife and with my child.”

### Group 4 – Swimming Lesson (local parent spectated)

In Group 4, a mother took her two sons (ages 8 and 13) to their swim lessons while a family friend (Jenny) connected from home via a laptop. Due to the noise in the pool area, the mother used a headset to hear Jenny better. She used the iPad first and then switched to the E2G prototype.



Figure 7. Swimming lesson with remote friend watching.

With the iPad, she stood at the edge of the pool to show the children swimming, primarily using the rear-facing camera. She had to be careful about getting splashed from the pool. She also wanted the children and Jenny to see each other, so she turned the iPad screen to the kids but struggled with displaying the correct (front-facing) camera.

When she switched to the E2G prototype, the mother sat down on a nearby bench to manipulate the camera. She was able to capture the children at any position in the pool while maintaining her conversation with Jenny.

During the interviews, both participants indicated that the interaction was very natural, especially when the remote person could see a zoomed in view of the event and the local person at the same time. Jenny said that *“it felt like I was sitting next to her when I could see her reactions and the kids swimming at the same time”*. However, Jenny also reported that the view was choppy when it was zoomed in. This could be because the zoomed in view was disorienting to her when the mother panned and tilted the camera.

When asked what she liked best about the experience the mother explained *“it was wonderful to share my kids’ activities real-time rather than just shooting video and emailing it later. It was fun to do.”* She also commented that she liked *“the immediacy of reaction with Jenny and her ability to ask questions and get answers real-time.”* Jenny also suggested improvements related to multiple views: *“I enjoyed the picture-in-picture but sometimes I wanted to be able to move it so I could see what it was covering up... or to turn that part off for a moment so I could see the whole picture of what the kids were doing.”*

### Group 5 – Swim Team Practice (local parent spectated)

The Group 5 session involved a father watching his 12 year old daughter at her swim practice at a public pool, while the mother connected from home. Because of the noise in the pool area, the father had to use a headset to talk with his wife. He started with the iPad and primarily held it by balancing it on his knee. He periodically switched between the front- and rear-facing cameras on the iPad as he attempted to capture their daughter swimming but also engage in a conversation with his wife.



Figure 8. Swim team practice with remote Mom watching.

Next, he used the E2G prototype. He controlled the camera and zoomed into various shots of their daughter without difficulty. He and his wife conversed the whole time.

During the interview the father commented that he and his wife talked both about their daughter’s swimming and general things related to their daily life. Both parents indicated that it was important to be able to see the other person as well as their daughter’s swimming, and for this reason they liked the E2G system.

When asked what she liked best about the experience the remote mother commented: *“I liked being able to talk to my*

*husband and also be able to see what was going on with my daughter.”* The father indicated that he liked being able to *“talk and show the event”*. He also commented: *“the iPad is handy to just grab and go, making it portable. The camera + tripod takes a little more planning but would be great to capture special events where a loved one can’t be there.”*

### Group 6 – Gymnastics (local parent spectated)

For Group 6, a mother took her daughters (ages 7 and 9) to a gymnastics class while the father connected from home. The father was attending to both the gymnastics lesson and the dinner he was cooking. The mother wore a headset to hear the father over the noise of the gym. While she appeared to know several other parents at the gym, she interacted with them only once.



Figure 9. Gymnastics class with remote Dad.

They started with the E2G prototype and then used the iPad. With the E2G prototype, the mother sat on the benches and used the camcorder to zoom in, across the gym. With the iPad, she switched to the rear-camera and walked around the gym to get close to where her daughter was practicing. At times, she even removed her shoes and walked across the mats to capture a better shot. Throughout the session, the mother and father talked continuously. The video of the father was oversaturated because of a bright window behind him, making it difficult to see him. Also, due to technical problems, the iPad was unable to receive the father’s video at times (the audio still worked fine).

When asked what they liked best about the experience, the mother commented that it was *“very cool to show my husband what the girls do at gymnastics. It felt like we were together.”* The father commented: *“I enjoyed the opportunity to watch my children’s event. Being able to talk with them and see their event while preparing dinner from home was awesome.”* The mother also explained that she *“liked the clarity of the camera + tripod, but it limited motion”*, and she *“liked the mobility of the iPad but it did not allow zoom/close-ups.”* The remote father felt that it was important to see both his wife as well as his daughters’ activities, and hence he preferred the E2G system.

### Group 7 – Outdoor Soccer (local parent spectated)

The Group 7 session involved a mother taking her 4-year old son to his soccer practice while the father connected from home. They first used the E2G prototype and then the iPad. Because the practice was at an outdoor field, the mother chose to have the tripod raised up high so she could stand while both watching her son and talking to her husband. The

## Computer Supported Young People

mother would also occasionally talk to her son or the coach during the practice.

While using the E2G prototype, the mother frequently adjusted and zoomed the camera to keep their son in frame. During one water break, the son came over and wanted to talk to his father. When using the iPad, the mother primarily used the rear-facing camera and walked up and down the sideline, trying to keep her son in frame. In general, the parents talked throughout most of the session, often talking about their son's behavior during the soccer practice.



Figure 10. Outdoor soccer practice with remote Dad.

The mother felt that optical zoom was important because she could focus the camera on their son and talk to her husband. She preferred the simultaneous dual views of the E2G prototype "because I was able to interact with my husband at the same time as he was watching. With the iPad you either had to be flipping back and forth or you could only see one thing." Both parents felt that seeing each other was important as well as being able to see their son's soccer practice. Hence, both preferred the E2G prototype.

### Group 8 – Indoor Soccer (local parent spectated)

For Group 8, the mother took her daughter (age 6) to her indoor soccer practice, while the father connected from his office. The mother used a headset to hear over the noise in the arena. They first used the E2G prototype followed by the iPad. The arrangement of the slate on the tripod made it difficult for the mother to sit, control the camera, and be face to face with her husband. Nevertheless, she felt that it was simply "kind of a new thing that you've gotta get used to following someone around." During a water break, the daughter came over to her mother and her mother tried to get the daughter to say hello to her father.



Figure 11. Indoor soccer practice with remote Dad.

With the iPad, the mother put a lot of effort into capturing a good view for her husband, walking everywhere to try to get close to her daughter. The mother stood to use the device but seemed to have to awkwardly look around the iPad to see what was going on for herself. "I generally liked the freedom

## February 23–27, 2013, San Antonio, Texas, USA

of moving the iPad around. I didn't like that you couldn't zoom in... I liked the freedom of being able to move around more and then switch from front to back."

Overall, this couple spent very little time talking, and primarily used the session as an opportunity for the father to watch his daughter. While the mother said it was somewhat important for her to see her husband, the father did not feel that it was important to see his wife's video.

### Group 9 – Track Event (local parent spectated)

In Group 9, the father attended a district high school track and field event with his daughter (age 16) while the mother attended from home. The father wore a headset so that he could hear over the field announcements.



Figure 12. Track and field race with remote Mom.

The father was very comfortable using the camcorder mounted on the tripod. He first showed his wife the field and tried to find other kids they knew. He then zoomed in to where his daughter waited with her friends for her event. When the race started, he panned the camera to follow the event. Interestingly, the starting line for one race was on the other side of the track, which was difficult to see from the stands. The camcorder not only provided the remote mother with a view of her daughter, it also enabled the father to see his daughter at the starting line. "What my wife liked is what she could see. It was like, 'Hey, I've never been able to see her get into her blocks'...being able to zoom way over there...she kept saying, 'Hey, this is kinda neat!'"

When using the iPad, the father primarily used the rear-facing camera, even after his daughter's event was over and he returned to conversing with his wife. However, the mother was not able to see as much detail as she could with the E2G prototype.

The father commented that he liked being able to "interact and share the activity". He was very comfortable with the tripod and liked being able to zoom in with the E2G prototype. Unfortunately, the mother did not submit her questionnaire so we do not have her feedback.

### Overall Preference

The goal of this study was not to compare the iPad to the E2G prototype, but instead look at the participants' overall preferences to provide insights into the importance of the features of each device. Of the seven groups that tried both devices, 12 participants expressed a preference (two remote participants did not complete these questions). In what follows, we represent a participant by their group number followed by L for local and R for remote person.

Looking at the absolute preferences, six participants preferred the iPad (3L, 4L, 6L, 8L, 4R, 8R). Preferring the iPad correlated with the importance of mobility for the event and not being concerned about seeing or socializing with the other adult. All three participants who expressed that they did not care whether they could see their partner favored the iPad (6L, 4R, 8R) as did the couple who did not socialize much and instead focused primarily on allowing the remote father to see the activity (8L, 8R).

The remaining six participants preferred the E2G prototype. Remote participants liked seeing both the local adult and the activity at the same time (5R, 3R, 6R), and local participants used the prototype in situations where mobility was either not critical or not possible (5L, 7L, 9L).

**Engagement**

One of the key goals of this study was to better understand how to support a feeling of *togetherness*, where families felt like they were experiencing the event together. The questionnaire posed three questions related to this aspect, “Compared to other forms of videoconferencing:”

1. it felt more like my partner and I were interacting naturally, as if we were together
2. it felt more like I was experiencing the activity with my partner
3. this experience was more engaging

Participants were asked to provide ratings for each of these questions on a seven point scale (1=strongly disagree; 7=strongly agree). For the 12 participants who provided ratings for both the E2G prototype and the iPad, we analyzed the responses using Wilcoxon Signed Ranks tests. Table 1 shows the mean and median results for each question, along with the number of people who gave each prototype the higher rating.

	E2G	iPad
Q1 # who gave higher rating	7	1
mean (SD)	5.3 (1.2)	4.5 (1.6)
Q2 # who gave higher rating	8	1
mean (SD)*	6.0 (1.2)	4.8 (1.7)
Q3 # who gave higher rating	7	2
mean (SD)*	6.1 (1.0)	5.1 (1.6)

**Table 1. Engagement ratings for E2G and iPad (\*p<.05).**

For all three questions, more participants rated the E2G prototype higher. For both Q2 and Q3, the difference in ratings was significantly different,  $Z=-2.35$ ,  $p=.019$  and  $Z=-2.0$ ,  $p=.046$  respectively, while the difference for Q1 was not statistically different,  $Z=-1.653$ ,  $p=.098$ .

These results suggest the combined ability to zoom in on the action and for the local and remote person to see each other is important for the feeling of togetherness. The participants’ comments offer further evidence of this. Several participants explicitly commented that they enjoyed the experience and felt very connected to their partner. 6L exclaimed: “*It felt like we were together. Personally, I loved the engagement I felt with my husband.*”

*Very cool!!”* 4R reported: “*I felt like I was sitting next to her when I could see her reactions and the kids swimming at the same time.*” 5R also commented that the shared context that multiple videos provided meant that “*I could sit and talk like I was next to him*”.

**Technology concerns**

Several key concerns related to both the E2G prototype and the iPad were observed during the study.

*Hands-Free Mobility*

Mobility is important for the types of scenarios examined in this study. First, all of these activities took place outside the home, so any technology being used needs to be transported to the event. This was problematic for our current E2G prototype since it is relatively large and awkward. Several participants commented that this was the key drawback of the device. Second, if the local adult needed to move around during the activity (as was the case for the play sessions), mobility became even more critical. Third, holding a device in-hand for an extended period of time is inconvenient if the adult is participating in the activity and tiresome over long periods of time. Ideally, the local adult should have a device that is both mobile and hands free.

*Video*

The E2G prototype was better able to capture most of the kids’ activities because it utilized a camcorder with 30X optical zoom. The lack of camera zoom was the biggest concern for most people using the iPad (7L: “*[I] liked the mobility of the iPad but it did not allow zoom/close-ups*”). The video from the iPad was reasonable when the local adult could get close enough to the activity, but the remote adult had a hard time following far away activities.

Unlike the iPad, the E2G prototype was able to simultaneously show the view of the local adult and the activity, and many of our participants expressed a strong desire for this feature. E2G shows both video streams to the remote adult using a picture-in-picture configuration. However, several remote adults felt that sometimes the picture-in-picture view obscured key areas of the video, so they wanted the ability to move and occasionally turn off the inset picture.

With both the E2G prototype and the iPad, the local adult framed the view for the remote adult. While this makes sense for coarse camera positioning, continual fine-grain adjustments were tedious. Several remote adults commented that they would like to be able to pan and zoom the camera themselves (3R: “*I also would have liked to be in control of the zoom in and zoom out feature*”).

*Audio*

The audio in our experiment was problematic in most of the sessions. It was hard for the local and remote adult to hear each other given the ambient noise in the locations. As a result, the local adult in all but two of the sessions used a headset to improve audio quality. While the headset was beneficial in this regard, it also meant that only one person could talk to the remote adult, and the local adult was cut-off

from conversations with other local people at the event. Furthermore, the headset prevented the remote adult from interacting with other local people (including the kids).

### *Framing*

Keeping the children's activity in the camera frame was a challenge. In most cases, the E2G prototype provided better framing, but required the local adult to continuously adjust the camera pan, tilt, and zoom settings.

Another challenge related to framing was getting a reasonable view of the local adult. The positioning of the slate on the tripod often made it difficult for the slate camera to adequately capture the local adult, especially while they were operating the camcorder on the tripod.

Interestingly, in several of our sessions the children wanted to see the remote adult as well. For the E2G prototype, this worked if the child came around to the back side of the prototype (in front of the slate), but if the child interacted from the front of the prototype, they were able to talk to the remote adult, but could not actually see them. The same was true for the iPad, unless the local adult turned the iPad around and switched to the front-facing camera view.

## DISCUSSION

The results of this study have several implications for the design of future video communication systems.

### **Going Beyond Talking Heads**

At the highest level, our study shows that people want to share experiences outside the home with remote family and friends. These experiences go beyond talking heads on screens to configurations that enable remote users to feel like they are actually attending an event. Enhancing users' sense of presence is challenging, but as shown in our study, providing multiple camera views can help a remote user feel like they are watching the event *and* socializing with other people who are local to the event.

### **Designing for Multiple Stakeholders**

The design of systems for shared mobile experiences must consider the needs of all stakeholders. When Grandma is remotely watching her grandson's baseball game, it is easy to envision Grandma as the main stakeholder (as depicted in our opening vignette). However, our observations identified four distinct types of stakeholders, and it is important to design for all of them.

One stakeholder is of course the remote observer, such as a parent, grandparent, or close friend, who benefits by being able to observe the local event live. The primary need of the remote observer is to see and hear the event. However, as we observed in this study, enabling the remote person to also socialize with other stakeholders is valuable.

A second stakeholder is the local observer who is sharing the local experience with the remote participant. The local observer has two key roles: streaming the activity to the remote observer and socializing with the remote observer. We observed many pairs talking and interacting with each other during the event. While some of this interaction was

about the shared activity, it was often general discourse about daily life. This "small talk" (also found in [10]) is an important feature when people share activities in person, and mobile video extends opportunities for keeping socially in touch with remote people. For the local person, being able to socialize with a remote person can also make the activity more enjoyable by filling in dead time in the activity (e.g., time between kid's races, breaks).

A third set of stakeholders is the local participants engaged in the activity, such as the children in our study. Enabling remote people to attend the event means that it can feel like more people are joining to watch their activity. For instance, kids will feel like the local *and* remote adults are actually at their event, which is important to them. Many of the children in our study loved the fact that their remote family members could be at their activity, and we observed several kids waving to or trying to talk to the remote adult.

The fourth stakeholder is the social context around the shared activity. This social context is made up of the three sets of stakeholders already discussed and other people around the local observer (i.e., other people watching the activity). We observed that these people sometimes interacted with the local and remote observers. Such interactions are important for enhancing the remote attendee's feeling of being at the event. Also, the experience should not detract from the local participants' social interactions with others at the event. Therefore, it is important to consider the social context around the activity when designing systems to support shared activities.

Our findings complement those from studies of mobile video calls out in public [10], which identified issues around managing boundaries for privacy and distraction between the private video call and others in the surrounding area. Our goal was to *include* a remote person in the shared activity, enabled in part by the larger video window in either E2G or the iPad. Thus, our design implications include how to integrate the surrounding social context as a stakeholder, rather than maintain boundaries from it.

### **Supporting Diverse Scenarios**

Our participants illustrated a diverse range of scenarios for sharing experiences in mobile environments. The variety of these events and the different needs of the stakeholders identify several important considerations when designing technology to support these activities.

One key consideration is the wide variation of motivations and goals for joining in the activity exhibited across our nine groups. One of the more compelling scenarios is including geographically remote participants who could not otherwise attend, like the grandparents in Group 1. This distance naturally provides motivation for enabling opportunities to participate in kids' events and socialize more with the local adult family members. However, it is interesting to reflect on the motivations of the remaining eight groups that chose to involve a spouse or a friend that was geographically local. These "nearby" participants were involved in other activities

(cooking, working from home) or in settings that did not offer scheduling flexibility (working at the office). Remotely joining offered them more flexibility in terms of scheduling (avoiding travel time). Specifically, while doing other activities they could pay partial attention to the kids' event and socialize with adults at the event. While we did not initially envision these scenarios, our experiences suggest that if sharing mobile experiences becomes easy and satisfying, connecting nearby participants may become a more common scenario.

Another key factor is the degree of involvement of the local observer in the event. On one end of the spectrum, the local observer may only be a spectator, such as the father who watched swim practice from the pool side. At the other end of the spectrum, the local observer may be highly embedded in the activity, such as the father who helped his kids catch fish at the fish hatchery birthday party.

When the local observer is spectating, seeing the remote person may be important so that socializing is more natural and enjoyable. However, when the local observer is participating in the activity, it may be less important to see the remote person. In our study, we observed two activities where this requirement differed. In the fish hatchery scenario, the father divided his attention between helping his children fish and talking with his parents. Having a view of his parents was important when he or his kids were talking with them, but it was less critical when he was busy with his children. In contrast, during the play center scenario, the father was actively engaged in playing with his daughter and did not want to visually attend to the device. He also did not feel it was important to see his wife; therefore, an audio-only link may have been sufficient.

Another consideration is whether the local person is primarily stationary or moving around at the event. When moving around, it is important that the device is easy to carry or reposition. For example, the mother watching her daughter's soccer practice wanted to walk around to get a good view, which the small form factor of the iPad afforded. A counter example is the father watching his daughter's swimming lesson. He was primarily stationary, so he liked the hands-free nature of the E2G prototype, even though it was less portable than the iPad.

Another factor is whether the local person is close to or far away from the event. At one end of the spectrum, the local observer may be close to the event, or able to walk up close to the event, as with the mother at the gym with her children. On the other end of the spectrum, the local observer may be far from the event or physically prevented from coming close, as with the father watching his daughter's track and field event from the stands. When the parent is close to the activity, a typical webcam may be sufficient. However, when the parent is far away, a camera that supports optical zoom will likely be required, and may benefit both the local and the remote participant.

A further distinction is whether the people taking part in the activity (e.g., the children) or other local observers want to interact with the remote person. For example, the children in the fish hatchery session wanted to talk to their grandmother and show her the fish they caught. In order to support these auxiliary social exchanges it is important that everyone can see and talk to the remote person.

### **Providing Technological Support**

We used the iPad and E2G to probe two points in a wide design space of possible systems to support mobile shared experiences. By observing the usage of two devices with different capabilities and limitations, we gained insight and understanding into the important capabilities for enabling mobile shared experiences. As our observations showed, neither the iPad nor the E2G prototype was optimal for all the activities we observed. In this section, we discuss some of the design implications open for future exploration.

### *Mobility*

When the local participant is in a crowded space or wants to move around, having a smaller, more portable device, such as an iPad, is beneficial. At the same time, if the participant is involved in the activity, having to hold the device can be problematic. As we observed in Group 2, it was extremely difficult for the mother to hold the device while playing with her kids in the park, and it was not easy to set the iPad down and capture the event. Additionally, continually holding the device for long activities will get tiresome. More design exploration is needed for a compact yet hands-free devices, perhaps something that is wearable.

### *Pan-Tilt-Zoom*

When the local participant is far from the event, having optical zoom capabilities is important. However, when high optical zoom is used, such as the 30X in our prototype, it is important to stabilize the video or else the remote experience will be unstable and jerky.

Who controls the pan-tilt-zoom is also important. In our study, the local observer controlled the pan-tilt-zoom, but several people commented that the remote observer should control the camera. This feature would provide the remote person with the freedom to choose the view they were most interested in and free up the local observer to enjoy the event and to socialize with others. These observations suggest that future systems should include powerful pan-tilt-zoom capabilities that can be controlled by both local and remote participants.

### *Seeing Both People and Activity Simultaneously*

Many of the remote observers in our study expressed a strong desire to see both the local observer and the activity at the same time. The ability to see and hear each other can dramatically improve the social aspects of the experience. When providing remote observers with multiple video streams, it is also important to give them control over how each video appears on-screen. In the E2G prototype, the video of the local observer was fixed in size and inset into the video of the activity. Several remote adults expressed that

sometimes the picture-in-picture view obscured key areas of the video. As a result, future systems should allow flexible positioning of concurrent multiple views.

### *Robust Audio*

Ambient noise was often an issue in the shared experiences we observed. Using headsets can improve the quality of the audio between the local and remote person, but limits interaction with others. The local participants cannot easily talk to other local people at the event and interaction with the remote person is reduced since only the person with the headset can talk to them. A better solution may be to use microphones with ambient noise suppression and loud speakers. Of course, use of loud speakers may not be socially acceptable when there are strangers around. As is often the case with video communication, clear audio with appropriate volume control is crucial for a good experience.

### **CONCLUSION**

Our goal was to gain a better understanding of the types of experiences outside the home that people want to share and the pain points they encounter when using existing technologies to accomplish this. We carried out a field study using two technology probes with distinctly different capabilities as a way to get a wider range of insights into these scenarios.

Our study produced several insights. First, the results show that people want to share experiences outside the home with remote and nearby family and friends. While they understand the limitations of existing technologies, they are willing to struggle through them to share the experience with a remote loved one. Second, besides the remote participant, there is a rich social context of other stakeholders that needs to be considered in designing mobile shared experiences. Third, there is a wide variety of sharing scenarios with different participant goals and requirements. Finally, we propose design implications for future systems around mobility, zoom, camera control, video views, and audio.

In the future, we plan to expand shared experiences beyond those involving kids at a particular location to events with a broader range such as a hike or a visit to a zoo. We will also continue to investigate whether audio between the local and remote observers is sufficient in certain circumstances if they have a shared view of the event. Finally, we plan to explore additional technology probes that combine the mobility and sleekness of the iPad with the zoom and multi video streaming capability of the E2G prototype.

### **ACKNOWLEDGEMENTS**

We would like to thank our study participants for their time in sharing their experiences with us. In addition, we are grateful to Tom Blank and Kris Crews for helping us construct the E2G prototype.

### **REFERENCES**

1. Bly, S., Harrison, S., Irwin, S. Media spaces: bringing people together in a video, audio, and computing environment. *CACM* 36, 1 (1993), 28-46.
2. Brubaker, J, Venolia, G., Tang, J. Focusing on shared experiences: Moving beyond the camera in video communication. *Proc. DIS 2012* (2012), 96-105.
3. Gaver, W., Sellen, A., Heath, C., Luff, P. One is not enough: multiple views in a media space. *Proc. CHI 1993*, 335-341.
4. Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., Eiderbäck, B. Technology probes: Inspiring design for and with families. *Proc. CHI 2003*, 17-24.
5. Judge, T., Neustaedter, C. Sharing conversation and sharing life: video conferencing in the home. *Proc. CHI 2010*, 655-658.
6. Juhlin, O., Engström, A., Reponen, E. Mobile broadcasting: the whats and hows of live video as a social medium. *Proc. MobileHCI 2010*, 35-44.
7. Kirk, D., Sellen, A., Cao, X. Home video communication: mediating 'closeness'. *Proc. CSCW 2010*, 135-144.
8. Mueller, F., Vetere, F., Gibbs, M., Edge, D., Agramanolis, S., Sheridan, J., Heer, J. Balancing exertion experiences. *Proc. CHI 2012*, 1853-1862.
9. Neustaedter, C., Judge, T. Peek-A-Boo: The Design of a Mobile Family Media Space. *Ext. Abstracts Ubicomp 2010*, 449-450.
10. O'Hara, K., Black, A., Lipson, M. Everyday Practices with Mobile Video Telephony. *Proc. CHI 2006*, 871-880.
11. Raffle, H., Revelle, G., Mori, K., Ballagas, R., Buza, K., Horii, H., Kaye, J., Cook, K., Freed, N., Go, J., Spasojevic, M. Hello, is grandma there? let's read! StoryVisit: family video chat and connected e-books. *Proc. CHI 2011*, 1195-1204.
12. Romero, N., Markopoulos, P., van Baren, J., de Ruyter, B., IJsselsteijn, W., Farshchian, B. Connecting the family with awareness systems. *Personal and Ubiquitous Computing* 11, 4 (2007), 299-312.
13. Strauss, A., Corbin, J. Basics of Qualitative Research, Sage Publications (1998).
14. Vutborg, R., Kjeldskov, J., Pedell, S., Vetere, F. Family storytelling for grandparents and grandchildren living apart. *Proc. NordiCHI 2010*, 531-540.
15. Weisz, J., Kiesler, S., Zhang, H., Ren, Y., Kraut, R., Konstan, J. Watching together: integrating text chat with video. *Proc. CHI 2007*, 877-886.
16. Yarosh, S., Inkpen, K., Brush, A.J. Video playdate: Toward free play across distance. *Proc. CHI 2010*, 1251-1260.