Capturing Mobile Telepresence through Logging and Video Shadowing: A Two-phase Study Design

Ellen Isaacs¹ and Margaret H. Szymanski²

Abstract
We used a two-phase study design that incorporated a novel video shadowing approach to learn how close-knit groups use technology to stay connected while mobile and to identify their unmet needs. In the first phase, we conducted a logging study with small groups in which each member logged all their interactions over 3 days, enabling us to identify suitable groups for the second, video shadowing phase. In the second phase, we separately observed and video recorded each member of the selected groups over the same half-day period as they connected in both mobile and stationary settings, a novel approach that others have been reluctant to try. This targeted shadowing approach enabled us to gain a rich understanding of the groups’ interactions across multiple media and devices that would not be possible using other indirect methods.

¹ Palo Alto Research Center Inc. (PARC), Palo Alto, CA, USA
² Xerox Innovation Group, Webster, NY, USA

Corresponding Author:
Ellen Isaacs, Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304, USA.
Email: ellen.isaacs@parc.com
Keywords
mobile communication, video shadowing, ethnography, user studies, channel blending

This article overviews the methods we used to conduct a study for a client investigating people’s need to stay connected while mobile, which we refer to as mobile telepresence. Our study was unusual in three ways. First, we combined two study techniques: logging and video shadowing. Second, we used video shadowing in a variety of mobile settings. Third, rather than shadowing a single person or activity, we separately observed each member of small, close-knit groups on the same day as they went about their activities, allowing us to capture each person’s context and point of view as he or she connected and disconnected over time.

We began our study of mobile telepresence with the understanding that people maintain ongoing, intermittent “conversations” over time with their close contacts, using multiple technologies in different contexts (Licoppe 2004). We designed the study to address the question: How do people in close relationships manage the threads of their connected lives across time, space, and technologies? Our goals in conducting the study were to:

1. use different ethnographic methods to deeply understand how people currently use technology to stay connected while mobile and stationary; and
2. uncover small groups’ underlying and potentially unmet needs with regard to staying connected.

Ideally, our findings would enable us to generate novel ideas for future technologies to support mobile telepresence and help us understand how to design such technology to better support the way people stay connected in real-world situations.

Background

This study was initiated by a client who wanted to discover new product opportunities in the domain of mobile telepresence while also learning about our ethnographic methodologies, hoping to incorporate them into their product development process. To support this learning, one member of the client’s team joined four of our researchers to create a five-person core team with occasional participation from others at both companies.
We proposed studying mobile telepresence practices through video shadowing, a method that consists of observing and video recording people as they go about their ordinary activities (Wolcott 2010). Specifically, we proposed selecting several small groups of friends and simultaneously observing each member during the same day to see how, when, and why they interacted (or not) while mobile and stationary. Video shadowing is a time-consuming activity, both in collecting and in analyzing the data. Since the client gave us just 15 weeks to complete the study, we needed to carefully select groups that would generate rich video data.

Most shadowing studies screen specific locations at which a behavior of interest occurs. In this case, we had to choose people, making it difficult to scout the object of study ahead of time. To address this challenge, we added a first phase to the study, a logging phase that would let us preview potential groups’ level and type of interactivity.

In this first phase, we recruited a larger number of groups and asked each member to log their interactions over a 3-day period. These logs gave us a relatively detailed picture of how, when, where, and how often these groups interacted with each other (and others). We used the logs to carefully select candidates for the shadowing phase and to determine when to observe them (see Figure 1). In addition, we returned to the logs after the shadowing phase to evaluate whether our findings were supported by the data from the larger phase 1 sample.

In the literature, most studies of mobility and connectedness have relied on second-hand data. Although first-hand in-situ observation is generally seen as preferable, Grinter and Eldridge (2001) noted that “direct observation [is] highly impractical” (pp. 222–23), so they studied teenage texting through surveys, logging, and interviews. Ito and Okabe (2005) used similar methods to study mobile communication, noting how “notoriously difficult” (p. 258) it is to get first-hand accounts. Similarly, Ling (2004) used interviews and Katz (1999) used surveys to study people’s cell phone use. Some studies have analyzed first-hand data by recording voice- or text-based conversations (Isaacs et al. 2002; Laursen 2006; Licoppe 2004), but we know of no studies that have used video shadowing to directly observe how small social groups communicate while mobile.

**Phase 1: Logging Study**

The logging study was designed to collect detailed data on remote and face-to-face interaction patterns from 10 groups of close friends or family, with the aim of selecting the best candidates for the video shadowing and
grounding our analysis of the video data. Specifically, we asked participants to record, for three consecutive days, all their interactions with other members of their group as well as others they liked to stay in touch with. They did not record interactions with acquaintances or strangers. We included at least one weekday and one weekend day because we expected those activity patterns to vary and we wanted to select the best type of day for shadowing.

To understand people’s communication patterns, we needed to know how they interacted with people in their network. So participants logged each interaction by indicating:

- with whom they interacted
- when they interacted
- their location
- what technology they used
- the purpose of the interaction

In addition, people logged times when they would have liked to interact with someone but did not or could not, the reason they wanted to connect, and why they did not. We included this “desired” case to learn when their connection needs were not being met and why.

Each group member logged their interactions separately. At the end of each day, each person e-mailed us their written log and left a 5- to 10-minute verbal account of those entries on our voice mail, which gave us a more detailed understanding of the interactions.
We gave participants the choice of two logging forms for recording their interactions. One was a spreadsheet with columns indicating media types (including face-to-face) and rows indicating time (see Figure 1). When participants had an interaction, they selected the box for that time and media type and then entered their location, the person they interacted with, and topic of discussion.

The second logging form was a simple text document that provided the following prompts:

- Time:
- Location:
- Media/Device:
- Who:
- Purpose:

These prompts were repeated many times so the participant could sequentially record each interaction in a day.

**Piloting the Study**

Before initiating the study, we conducted a pilot study to test the procedure and materials. We realized that asking people to log all their interactions could be burdensome, so we wanted to see whether people would do so with enough frequency and detail to be useful. We ran pilot tests with two pairs, each for one day of logging and a half hour of video shadowing. We selected employees of our organization whom we knew to be highly connected.

We initially gave the pilot participants just the spreadsheet log form. Although they filled it out as instructed, they reported difficulties doing so while mobile. So we created the simpler text form and revised our instructions to allow people to track their interactions using any other method (e.g., taking notes on their smartphone) as long as they provided all the information required for each interaction.

Several days after each pair logged their interactions, we conducted a short video shadowing session, following each member of the pair for the same 30-minute period when they were both available for interactions. Even though we have done quite a bit of video shadowing, we still found the pilot helpful for learning how to capture small device screens as effectively and unobtrusively as possible as people moved around. It was also useful to show the pilot video data to our client during a mid-project workshop, as it helped them better understand our shadowing plans.
Recruiting

Our goal for selecting participants for the logging (and shadowing) study was to find people who (1) connect to many people, (2) connect frequently, (3) are frequently mobile, (4) use a variety of technologies to connect, and (5) have a range of attitudes toward technology. We were looking for people who are on the forefront of adopting mobile telepresence technologies and are developing practices for such activities. Our client also wanted to include some groups that were cross-generational, and some distributed across time zones. We used purposive sampling (Patton 1990) to choose such participants. We decided to limit groups to no more than three people to make the observation process more manageable, knowing we would capture additional interactions as participants encountered others not in the study.

To find participants, we used our personal networks and those of other employees. The client asked us not to use an open recruiting method such as Craig’s List for fear of attracting people who simply wanted to participate in a paid study and might not respond truthfully to a screening survey. We considered using a recruiting agency, but this would have added extra time and given us less control in interacting with the participants during the logging study. This control was important because our study was more complex than those typically handled by recruiting agencies, which generally point participants to an online survey or arrange for them to show up at the study location.

We developed a screening survey to identify highly connected participants and administered it through SurveyMonkey.com, a website that lets users generate a survey, collect responses, and collate the data. Table 1 indicates the key survey questions and our criteria for selecting participants.

We looked for people who met as many criteria as possible, and most met all of them. However, the vast majority of our candidates (28 of 35) were women, probably because we recruited people who considered themselves “highly connected.” To make sure we included some men, we relaxed the criteria for number of connections per day and, in one case, the number of media types used.

When potential candidates contacted us, we sent them an e-mail describing the study requirements and procedure and explaining that they must to be willing to participate in both phases of the study even though only some would be selected for the shadowing phase. It directed them to the screening survey if they were interested in participating.

We received 48 e-mail inquiries from potential participants and 35 people filled out the screening survey, making up 14 potential groups.
From those, we selected 10 groups for the logging study—seven pairs and three triads consisting of 16 women and seven men. Two groups included participants split between time zones. One group was cross-generational and one consisted of two middle-age women. The rest were people in their late teens to early 30s. Each participant was paid $150 for the logging phase and $300 for the video shadowing phase (if selected).

**Human Subjects Approval**

Before running the study, we solicited human subjects’ approval from our company’s Internal Review Board (IRB), an accredited organization tasked with protecting the rights of study participants. Our application described the study design, our recruiting method, study compensation, our method for protecting the confidentiality of the data, and any risks to participants.

The IRB raised a concern about video recording bystanders, and we agreed to get written consent from people who substantially interacted with the video shadowing participants and verbal consent from those they interacted with briefly. We were not required to get consent from

---

**Table 1. Key Survey Questions and Criteria for Selecting Participants**

<table>
<thead>
<tr>
<th>Question</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>For each member of your group, which methods do you use to stay in touch with each other?</td>
<td>Should select at least 5 methods</td>
</tr>
<tr>
<td>About how often do you stay in touch with each person?</td>
<td>At least 1–2 times per day; if a triad one person should be 3 times per day or more</td>
</tr>
<tr>
<td>How close do you feel to each person?</td>
<td>At least one person should be rated “very close” or “extremely close/intimate”</td>
</tr>
<tr>
<td>About how many friends and family members do you connect with at least once on a typical day?</td>
<td>At least 4–6</td>
</tr>
<tr>
<td>Approximately how much time do you spend on average at the following locations?</td>
<td>At least 1–2 hours each “traveling around,” “walking around,” and/or “in other locations” (or a combination that added up to at least 4 hours)</td>
</tr>
<tr>
<td>Attitudinal questions about value of personal relationships and technology.</td>
<td>High value on personal relationships, and a variety of attitudes about technology</td>
</tr>
</tbody>
</table>
bystanders captured in the background. In practice, the participants informed their friends about the shadowing and arranged to meet with only those who were willing to be recorded, so written consent was agreed on ahead of time. When they spontaneously interacted with others, we either avoided filming them or got verbal permission when the person saw the camera. Designing the study, preparing the materials, and recruiting participants took 3½ weeks, including a client workshop.

Conducting the Logging Study

Each research team member was assigned one to three groups and was responsible for carrying out each step of the study with those groups. These steps consisted of contacting participants, running the setup interview, collecting the data, conducting the exit interview, and analyzing the data. To ensure that everyone followed the same procedure, we created a setup and interview guide and an exit interview guide, which laid out all the steps of the setup and debriefing process.

Participants were initially contacted via e-mail or phone. The researcher explained the procedure and arranged to meet for the setup interview. We e-mailed the participants the consent form ahead of time so they could read it carefully and return it during the interview. We also e-mailed them the logging forms and instructions, which we explained during the setup interviews.

The setup interviews were conducted face-to-face with the whole group at their home or another convenient location. Any remote members were included via conference call. During the interviews, we explained in detail how participants should record their interactions and give their daily verbal report, showing them a sample log and playing a sample verbal report. We also gave them written instructions with a short “cheat sheet” explaining the key aspects of the procedure. Next, participants chose a consecutive 3-day period to log their interactions that included at least one weekday and one weekend day.

Then we conducted a short pre-logging interview to get background information such as how and when they met, how often they got together, the communication media they preferred and why, and so on. We also asked them to describe their plans for the days they would be logging. The setup interviews lasted 30–60 minutes and were audio recorded.

Once logging began, each researcher reviewed their groups’ logs and verbal reports each day to make sure the participants were doing them properly. Overall, the participants complied with the instructions very well,
although we contacted a few participants to ask them to provide more or sometimes less detail in their remaining reports. Participants’ verbal reports usually lasted less than 10 minutes, but a few lasted over 20 minutes. While reviewing the data, we noted any questions or interesting situations to probe during the exit interview.

After each group finished logging, the assigned researcher met with them for an exit interview and debriefing, mostly in person but sometimes via conference call. During this interview, we asked a few basic questions about the study itself and how the logging procedure may have affected their interactions. We followed up with more in-depth probes about specific interactions reported in the logs. Finally, we answered participants’ questions, arranged payment, and discussed possible participation in the next phase of the study. The interviews lasted 30–60 minutes and were audio recorded.

**Analyzing the Logging Study Data**

Since our objectives for the logging study were to identify appropriate candidates for video shadowing and to get an overview of interaction patterns common among highly connected people, we carried out a high-level analysis of the data. Given more time, we would have mined these data more thoroughly, but the project timeline required us to be strategic in our analysis.

By the time we collected all the logs, it was apparent which groups were good candidates for video shadowing based on the number of connections reported, the variety of media used, and the richness and variety of those interactions. Although one might assume that the screening survey provided much of this information, our data confirmed our belief that when people summarize their behavior they are less accurate than when they record individual instances. We determined that 6 of the 10 groups would make good video shadowing candidates, and we selected four of them on the basis of scheduling issues, their consistency during the logging study, and our goal of including at least one group with a remote member and one that included men. The selected groups were two pairs and two triads made up of eight women and two men. As it turned out, none were personally known to the researchers. This filtering process reduced the risk that our video shadowing efforts would not yield a wide variety of interesting remote communication activities.

The logging data analysis revealed some common themes of the participants’ interactions and generated a basic quantitative characterization
of their interactions (number of interactions, plus number of conversations carried out over time, across media, with multiple participants, etc.). Each researcher took notes based on the verbal reports, written logs, and recordings of the setup interviews. Our notes captured any aspects of the interactions that were of interest for any reason. Each noted item was recorded as a bullet posted to an internal website so we could all see the data. Altogether, we generated around 250 descriptive notes from the 10 groups.

Using these notes, we did an affinity mapping exercise (Strauss and Corbin 1990) to pull out the themes from the data. Using this approach, all the bullet items from all the groups were combined in one big pile, and then sorted into groups based on their similarity; items could be placed in more than one group. This approach can be quite subjective, as it is up to the analyst to decide what is similar. The groups of related items were labeled to denote their theme. In this way, patterns hidden in the data emerged in a bottom-up process.

To assist us in understanding how people carried on threads of conversations with multiple people over different media, we experimented with ways to visualize the log data. We settled on the approach shown in Figure 2, in which we color coded the cells based on the person the participant interacted with, and then drew lines between cells where the same topic was discussed. This visualization allowed us to see when a topic was discussed across media and with different sets of people.

While this visualization helped us understand the data, it was not so effective for communicating the results to our clients. So we generated an alternate visualization, shown in Figure 3. Again, colors indicate people and lines show topic threads, but this visualization had more impact because it is more compact and visually compelling. At a glance, the viewer can tell whether a person had many or few conversation threads, whether those threads involved many or few people, whether they lasted for many or few interactions, and whether they crossed many or few media. For example, without fully understanding the notation, it is easy to see in Figure 3 that Subject 1a’s interactions were less complex than Subject 2a’s. These visualizations effectively gave our clients an intuitive sense of the data.

The logging phase of the study, including analysis and a client workshop, lasted 5 weeks and overlapped in part with the video shadowing phase.

**Phase 2: Video Shadowing**

In the video shadowing phase, we selected four groups and arranged to observe each group for a 4- to 6-hour period when everyone would be
available for interactions. In each session, each group member was accompanied by a different researcher who observed and video recorded them as they each went about their activities. The goal of this approach was to capture all sides of interactions and understand each person’s context leading up to, during, and following those interactions. As mentioned, we believed this approach to be unprecedented, particularly for a mobile technology study, so we were hopeful it would yield interesting and novel findings.

To set up the shadowing session, the assigned researcher worked with the participants to arrange a suitable time, taking into account the patterns revealed by their logs and their upcoming plans. In three cases, we met them after work and recorded throughout the evening and in the fourth, we recorded on a Saturday. In one case, a triad was split between California and Texas, so one researcher flew to Texas and we observed the group for the same 5-hour period, adjusting for the time difference. The remaining groups lived in the same region of California.

**Figure 2.** A portion of a participant’s log showing our visualization of her interactions. Each box indicates a single interaction and is shaded a color that indicates the person involved in the interaction; the lines connect interactions on the same topic.
Table 2 shows interactions occurring in four contexts.

To improve our chances of seeing examples of interactions in all four quadrants, we asked the participants to make plans to go out in public and, if they were local, get together at some point. Although the goal of video shadowing is to capture natural behaviors, we felt this guidance was appropriate since people would still be doing activities of their choosing and behaving normally during those activities.

The participants all wore wireless microphones to ensure quality audio recording. When possible, we equipped the cameras with wide-angle lenses to better capture interactions at short distance.

Although we had some concern about the difficulty of video recording people in public and while mobile, we were pleasantly surprised at how successfully we were able to capture an extensive and interesting variety of interactions. Only once was a researcher asked to stop recording (in a

**Figure 3.** Alternative visualization of two daily logs. The squares indicate people the participant interacted with at different times (with each person assigned a different color), and the lines show the same topic being discussed across interactions.
department store), but he later got management approval and was able to continue. Otherwise, we were able to record continuously across a variety of places, including restaurants, grocery stores, coffee shops, a dog park, a farmer’s market, a college campus, on busy streets, in cars, and in people’s homes. Although the participants knew they could ask us to stop recording at any time, no one did so.

In crowded public spaces, it worked best if we held the video camera at waist level and angled it up, thus attracting less attention. We could monitor the participant through the video display without having to look through the viewfinder. When a participant moved quickly through a crowd, it was helpful to follow closely and aim the camera at the participant’s back, allowing us to track her or him and preclude others from separating us. We found it helpful to wear a small camera bag across our body so we could rest the camera on the bag, which kept our arms from getting tired and stabilized the image.

We experienced an interesting challenge related to shadowing multiple people at the same time. Since each group member was accompanied by a different researcher, when they got together face-to-face in small spaces (such as an apartment), the room got crowded, which may have drawn their attention to the fact that they were being recorded. (In one case, a client researcher joined us for the shadowing to learn how it was done, which meant we had more researchers than participants!) Usually, we adjusted so that one person shot the overall scene and the other tried to capture the devices being used, but that varied depending on the circumstances.

Another challenge was avoiding becoming a member of their social group. In two instances when the shadowing sessions spanned lunchtime, participants invited the researcher to join them for lunch. In one case, this caused a breakdown in the observer–participant framework that was unrecoverable; the post-lunch interactions included the researcher as a fully ratified participant. Other than these cases, the participants generally became accustomed to ignoring us fairly quickly and only interacted with

| Table 2. The Contexts in which We Hoped to Collect Interactions during Video Shadowing |
|------------------------------------|-----------------------------------|
| **Static**                        | **Mobile**                       |
| Face-To-Face                      | For example, at a restaurant      |
|                                   | together, driving together        |
| Mediated                          | For example, phone call,          |
|                                   | video chat, Facebook              |
|                                   | For example, texting or phone call while mobile |
us on logistical matters (e.g., arranging seating in a car) or to clarify certain points (e.g., whom they were speaking with on a phone call).

**Analyzing the Video Shadowing Data**

Analyzing the video data was an iterative process that involved the core group, with occasional participation from the extended team. It consisted of clipping video segments of interesting interactions, transcribing them (Atkinson and Heritage 1984), looking at them as a group to understand the interactions, and in some cases searching for other similar interactions in other videos to create a collection of related instances. From the roughly 50 hours of video data collected, we identified more than 80 episodes of mobile telepresence using mobile phones (for voice or data), texting, e-mail, social networking sites, instant messaging, video conferencing, photo and video sharing, simultaneous web browsing, and landline phone calls, plus many face-to-face interactions involving the sharing of remote data through mobile devices.

After each shadowing event, each researcher viewed the video they had recorded and selected short segments (generally 1–4 minutes long) that captured interesting interactions. It was up to each researcher to decide what was interesting, given the goals of the study, although we usually discussed which clips to focus on after returning from a shadowing outing. We were particularly interested in analyzing the interactions involving all or some of the group’s members. We also met informally with one another to show each other potential segments to clip and to discuss why they were of interest. In some cases, we transcribed the videos right after clipping them; in others, we waited to transcribe until we had shown the clip to the group and everyone agreed it was worth analyzing further.

In cases where we had captured the same interaction from multiple participants’ point of view, we generated videos that combined each person’s video in separate portions of the screen and layered together the audio from each video stream. These “mosaic videos” allowed us to watch the interactions unfold from everyone’s point of view simultaneously. Combining multiple points of view was especially interesting when the participants were interacting remotely, but it was also useful when they were co-located because we could see the interaction from multiple perspectives. Since each participant was wearing a wireless microphone, having multiple videos enabled us to generate complete transcripts even when there was overlapping speech.

During the data sessions, we collaboratively analyzed the clipped videos using conversation analysis (Sacks 1984, 1992) to understand what was
happening sequentially from the participants’ point of view. In these structured data sessions, researchers repeatedly reviewed and discussed the details of particular sequences of activity; colleagues built on each others’ observations and insights were uncovered that no one saw initially. Each person brought this renewed perspective to their preliminary review of new data, helping them select appropriate segments for further review. In this way, the team built up collections of related interactions. Jordan and Henderson (1995:5) describe the foundations of interaction analysis and discuss the value of these types of group data sessions, which they call interaction analysis laboratory.

As we came to understand the data set, we discussed its significance to our research questions. As we identified patterns, we discussed whether we had seen any similar interactions in other shadowing sessions, perhaps involving different technology or in different settings. If so, the person who had recorded that interaction clipped the new video segment and we discussed it in a later data session. We repeated this process to build up our understanding of the participants’ behavior, the problems they encountered, and how they handled those problems. This process enabled us to generate a range of novel insights regarding mobile telepresence. For example, a key finding was that people frequently tried to integrate both local and remote participants into a single conversation, often trying to share digital or physical content with both sets of people, even though the technology did not support such activity or made it difficult. We named this phenomenon channel blending and we discuss it in detail in Isaacs et al. (2012).

In addition to using the video for our analysis, we used it during client workshops to communicate our findings and provide training on video-based interaction analysis.

As a final step, once we identified themes about people’s mobile telepresence practices, we returned to the log data to see whether they supported those findings. In doing so, we saw patterns that had escaped our notice during the initial logging data analysis and indeed validated that the themes appeared in many groups, not just the ones we shadowed. For example, we found that when channel blending, it was common for a “pivot person” to attempt to integrate the channels from local and remote participants to create one coherent conversation, and we identified new combinations of media being blended that we had not seen while video shadowing (Isaacs et al. 2012).

The video shadowing phase lasted 6½ weeks, including a client workshop, followed by an additional 2 weeks to write up the final report. This was a compressed time frame that gave us just enough time to generate.
preliminary results. We were fortunate to be able to spend another 3–4 weeks after the client engagement ended further analyzing the data to gain a deeper understanding of the phenomena we had uncovered. We recommend planning for this extended analysis period in future research.

Concluding Remarks

As ethnographic research becomes less tied to a particular place, mobile data collection techniques will continue to develop. We used a video shadowing technique usually used in stationary settings and applied it to a mobile setting, a challenging endeavor that others have been reluctant to undertake. We further explored a novel approach in which we simultaneously observed interactions from each group member’s point of view, often from different locations. Although it was initially daunting to consider shadowing people using small devices as they moved around, sometimes in public places, doing so turned out to be a manageable undertaking and was well worth the challenge. By following each person’s context and watching interactions as they unfolded, we achieved a deep level of understanding about the dynamics of those interactions that revealed insights (Isaacs et al. 2012) that would not have been available to us otherwise.

Video shadowing requires extensive, time-consuming data analysis but we had relatively little time, so we needed to choose groups that would yield rich and fruitful data. To do so, we preceded the study with a logging phase in which groups recorded specific information about their interactions for 3 days. These logs gave us a preliminary picture about when, where, and how the groups interacted so we could select suitable candidates and identify fruitful times to observe them. Later, the log data provided a means of validating the video shadowing findings with a broader sample and showed how the behavior evolved over days rather than hours. A phased approach to data collection is one way of ensuring that the time spent in the field yields valuable data and a successful research outcome.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.
References


