

Voice-Mail Diary Studies for Naturalistic Data Capture under Mobile Conditions

Leysia Palen

Department of Computer Science
University of Colorado, Boulder
Boulder, CO 80309-0430 USA
+1 303 492 3902
palen@cs.colorado.edu

Marilyn Salzman

Salzman Consulting
878 West Tamarisk Street
Louisville, CO 80027
+1 303 666 6472
marilyn.salzman@acm.org

ABSTRACT

Mobile technology requires new methods for studying its use under realistic conditions “in the field.” Reflexively, mobile technology also creates new opportunities for data collection while participants are remotely located. We report on our experiences with a variation on the paper-based diary study technique, which we extend by using voice-mail paired with mobile and landline telephony to more easily collect data in natural situations. We discuss lessons learned from experiences with voice-mail diary studies in two investigations of different scope. We also present suggestions for tailoring the technique to different research objectives, garnering high subject participation, and configuring the voice-mail system for data collection.

Keywords

Diary study, mobile computing, mobility, phones, usability, voice-mail, naturalistic study, HCI, CSCW

INTRODUCTION

We have extended the diary study technique to examine and leverage mobile devices in everyday use. Our development of voice-mail diary studies arose out of a need to understand how people use mobile technology across a variety of situations in everyday life. Paper diary studies have been used effectively for the collection of naturalistic data in human-computer interaction (HCI) and computer-supported cooperative work (CSCW) research [1, 2, 3, 4, 7, 8, 12, 14, 15, 16, 17, 18, 21]. However, the ecologically valid study of today’s mobile technologies is challenging because the conditions for technology use are dynamic, varied, and difficult for investigators to directly observe. Reflexively, mobile technology—most notably, mobile telephony—has also created opportunities for capturing everyday activity in new ways.

In our investigations of mobile telephony use, we extended the diary technique by using telephony and voice-mail. We have briefly written about our data collection approach elsewhere [16, 17] as a function of reporting on the entire research project; however, when presenting that material,

we have fielded many questions about our implementation of and experiences with voice-mail diary studies. These experiences and discussions with research scientists and usability specialists point to a need for efficient, non-intrusive data collection methods that yield informative, naturalistic data when studying issues of mobility.

We found the voice-mail diary technique to be very useful in this regard, especially when used in supplement to other data collection methods. Therefore, we offer this paper as a partial response, with an in-depth discussion of the voice-mail diary study technique (see Mark’s approach [11] as another interesting response). Our discussion is not intended to be a rigorous analysis of the diary study method in comparison to other data collection methods. Rather, this paper is meant to serve as an introduction to an approach that can be creatively extended to support many research objectives.

We begin with a review and survey of diary studies in HCI and CSCW research. Then, in the context of two voice-mail diary studies, we discuss the quality and quantity of the diary study data, the degree and nature of subject participation, and the lessons learned from the method’s different applications. We also present additional methodological concerns, including how implementation of a voice-mail diary study might vary when the mobile phone used for the diary is also the subject of study. We offer suggestions for participant compensation schedules, and adaptations of the method for different research objectives. Finally, we address practical voice-mail system implementation issues that would be of concern to investigators, including voice-mail set-up, transcription techniques, and opportunities for utilizing the medium for additional communication with participants.

DIARY STUDIES

Diary studies have roots in multiple disciplines. Psychology is the field that appears to have most influenced diary study adaptations for HCI research, but variations on the technique have histories in health and medicine, education, anthropology, and architecture, among others. More recently, diary studies have been adapted from HCI to CSCW investigations to further reflect a sociological interest in highly descriptive accounts of activity.

*To appear at Computer Supported Cooperative Work
Conference 2002 (CSCW 02). New Orleans, LA,
November 2002.*

What Are Diary Studies?

Diary studies, as used in HCI and CSCW research, are designed to capture activities that occur in real environments vis-à-vis some kind of technology currently under investigation, or one subject to design. Participants are asked to record particular activities as they occur on a paper diary. These diaries can be highly structured, with specific pre-defined categories of activities to be checked off and later counted, such as the number of incoming phone calls over the course of a work day [4]. They can also be unstructured, with spaces for recording, time-stamping, and describing activity (e.g., [1]). Activities of interest could be either seemingly mundane or exceptional; they could occur frequently or rarely. Ultimately, diary design is highly research-specific: Rieman, for example, was interested in “eureka” events with respect to photocopier use, and instructed participants to record sudden discoveries and insights as they occurred in everyday situations [18].

Implementation of diary studies often requires participants to frequently, and even daily, discuss diary entries with the investigator [18]. Frequent investigator involvement, especially initially, also ensures that participants understand the scope and depth needed for the diary entries. One of the challenges in diary studies, in fact, is convincing participants that seemingly mundane and low-level activities really are of research interest. Calibration of diary entries with investigators is necessary for garnering the type of data required. Another challenge that frequent investigator involvement helps mitigate is declining dedication to diary entry. Investigator involvement helps keep interest high and reminds participants of the importance of the diary in data collection.

Origins, Interpretations & Applications

Diary studies have methodological traditions in multiple fields, although the use of diary studies in HCI (and later, in CSCW) seems to be initially influenced by Psychology by way of Cognitive Science research. Norman and Sellen, for example, used diary studies in their human error work in the mid-80s and early 90s [13, 20]. Two publications from the early 1990s were the first to formally introduce the method to the HCI community, although others were also using it at that time for HCI-related research (including [3, 4]). Kirakowski and Corbett’s 1990 book on HCI methods [9] describes diary studies and even offers useful variations that we incorporated into our voice-mail method. Rieman’s frequently cited 1993 InterCHI paper [18] presents the diary study method as a middle-ground solution to the opposing limitations of laboratory studies and field studies. Diary studies can impose useful experimental constraints while maintaining ecological validity because they are conducted in natural settings but retain some level of researcher control [18].

CSCW researchers have made frequent use of the diary study method in a range of investigations on: reading practice [1]; the effects of technology failure on planned activity [7]; teenage short text messaging [8]; the temporal organization of computer-based authoring work [12]; information capture [2]; the information needs of mobile professionals [14]; and shopping behavior [15]. These

studies are illustrative of the anthropological and sociological influences on diary study approaches because they move beyond objectives of counting events to focus on descriptive accounts of activity. In particular, Brown, O’Hara and Sellen’s [2] and O’Hara and Perry’s [15] diary approaches echo the practice of Visual Anthropology to use photography as a research tool. In their adaptations, participants rather than the investigators took the photographs, resulting in a visual montage of activity across several participants and days. The investigators deployed cameras in lieu of paper to study participants, with follow-up interviews.

The Experimental Sampling Method (ESM) [10] is conceptually related to the diary study method. Developed in the mid-70s by Csikszentmihalyi, Larson and Prescott [5], the method is now used in a wide range of fields to systematically study subjective states of people in naturalistic settings [6]. Sometimes referred to as “beeper studies,” ESM studies require the distribution of pagers to subjects along with carefully designed questionnaires about their activities and psychological states. Investigators page subjects 7-10 times throughout the day (by design, but arbitrarily from the subjects’ point of view). Upon being paged, participants complete the questionnaires. In contrast to diary studies in which the subjects directly control when an entry is made, the recording of activities with ESM is investigator-determined.

ESM approaches have carried over into the practice of design in innovative ways. In particular, Rick Robinson, a former student of ESM innovator Csikszentmihalyi, adapted the method to workspace design at E-Labs [19]. Using cameras and pagers, occupants in a space subject to re-design simultaneously receive a page. They stop and immediately take pictures of their current activities and surroundings. The catalog of images provides a synchronized account of space use throughout the day from the perspectives of the users, but at times controlled by the investigators.

Voice-mail Diary Study Overview & Advantages

Our voice-mail diary study technique is a powerful though relatively simple adaptation of the paper diary study method. It extends Kirakowski and Corbett’s suggestion to incorporate audio recording [9], and adopts the spirit of Brown, O’Hara, and Sellen’s employment of photography [2] to capture naturalistic, in-the-moment experiences.

In the voice-mail method, participants use mobile or landline phones to make reports to a dedicated voice-mail line instead of recording events on paper. We found this to be an easy and less-time intensive way of reporting activities of research interest. With paper diaries, users must stop their activity and manually record it; this is not so troublesome when users work at a desk and the activities of interest occur there but becomes problematic when participants are mobile. Grinter and Eldridge [8] used a paper diary technique in their study of teenage users of short-text messaging (SMS), a text chat service available on mobile telephones. They recognize that having participants record SMS activity was problematic when it

occurred on buses, while walking along the street, or even in the bath [Grinter, personal communication]. Although using a mobile phone to call into a voice-mail line also means that subjects must stop their activity, they are often able to suspend and recommence activities more quickly, using a device they would likely carry anyway, and often providing richer description than if they were to make notes on paper.

Voice-mail diaries also have the advantage of giving the investigator immediate access to the reports without having to meet with participants face-to-face. Although frequent investigator involvement is an important part of a successful diary study, we found this did not necessarily have to happen face-to-face. Regular acknowledgement by the investigator of participants' entries and updated outgoing greeting recordings supported bi-directional communication. We discuss these issues in more detail in the remainder of this paper.

Variations on this technique can accommodate a range of research objectives and analytical styles. Like HCI-adapted paper-based diaries, voice-mail diaries can be structured or unstructured. In a highly unstructured diary, participants might provide open-ended, stream-of-consciousness narratives about their activities and experiences, which can result in very rich and detailed accounts¹. In a highly structured diary, participants might be supplied with a set of issues they report on by selecting numbered options in the voice-mail system. For the remainder of this paper, we describe our use of the method in detail and offer suggestions for variations on the method to suit different research objectives.

METHOD APPLICATION IN TWO STUDIES

We have used the voice-mail diary technique in two related studies, but adapted it to work with the studies' different objectives, scope, length, subject pool size, and available investigator time. The "Going Wireless Study," an open-ended, qualitative investigation, was our first use of the technique [16, 17]. We refined and adapted the method for use with the follow-on "Wireless Life-Cycle Panel Study." In both studies, participation in the diary portion of the data collection was voluntary though encouraged.

Study 1: Open-Ended, Qualitative Investigation

Description of the Going Wireless Study

The Going Wireless Study was an investigation of 19 novice users' adoption and use of mobile telephony over a

six-week period. Two investigators were on the project, with one investigator's time entirely dedicated to data collection and analysis.

We were interested in several issues: how usability affected use and discovery of handset features, how this in turn affected integration of mobile telephony into everyday life and the development of communicative practice, as well as how psycho-social factors affected the same. Since the study was broad in scope with the issue-space largely data-directed, we opted for a highly unstructured approach to diary data collection. In addition, we conducted three lengthy interviews with each participant and had access to calling data records.

Application of Method

Because we were concerned that requiring participants to use the method would erode participation in the six-week long study, we made participation in the voice-mail diary optional. In doing so, we regarded the data as supplemental to our other data sources, but hoped it would result in highly descriptive data we could not otherwise obtain.

However, we did create incentives to encourage participation. In addition to the \$100 participants received for participating in the three interviews, they also received \$1/day for every day they called in. Participants were given a list of issues to report on, including when they first used their mobile phones in a new environment or in an unexpected way, periods when they did not use their phones at all, when they used a new feature or had a problem. We asked them to report their contacts with the service provider. Again, our intention for the diary was not to collect rigorous, quantitative data; instead, we sought to learn about experiences that were significant to new users close to the time they occurred. We also wanted to develop a general understanding of the nature of use that could inform interview inquiry and the quantitative calling behavior data for each participant.

Additionally, because we were interested in studying the natural evolution of mobile telephony communicative practice, we never specifically instructed our participants to use their mobile phones to call us. Such instructions were likely to affect the use of our very object of study. Because we were interested in learning when in the adoption process people learned to program phone numbers into their phones, we never offered nor suggested that they program our voice-mail number into their phones. In fact, most participants called us from their home landline phones for some time before switching over to calling from their mobile phones. Some eventually programmed the voice-mail number into their phones, some continued to carry our business card with the voice-mail number, and one even taped the number to the back of her phone.

Method Results

Analytically, complementing our other data sources with unstructured voice-mail diary data was beneficial. At its most subjective, the data gave us a good feel for the nature of use of mobile phones for several of our participants. We had a sense for the kinds of activities that occurred in their daily lives. More objectively, we were able to use the

¹ A compelling though unempirical use of voice-mail was employed by the National Public Radio Sonic Memorial Project, which used it as a means to capture sounds that chronicle the "life and history of the World Trade Center neighborhood and working environment before, during and after" the attacks of September 11, 2001 [22]. Invited to call in to the dedicated voice-mail line, people offered songs, stories, previously recorded sounds from a normal work day in the towers, and even wails of grief. The recordings will be maintained as an audio memorial, broadcast on radio and available on the web [22].

information participants provided in their voice-mail reports to tailor interviews specifically to them. Furthermore, our long-term employment of the method helped develop positive researcher-participant relationships: It appeared to lend a sense of continuity to communications, even though the relationships were partially mediated by the voice-mail system.

Participation Rates. Because the diary study portion of the investigation was voluntary, participation rates varied from almost none to almost daily, although overall participation was higher than we had expected. The total participation rate across all participants (except S12) was about 42%, which we calculated as a ratio of the number of days diary entries were made by the total number of days in the study. Figure 1 presents the distribution across all participants (S12 left the study mid-way, so her data are not presented).

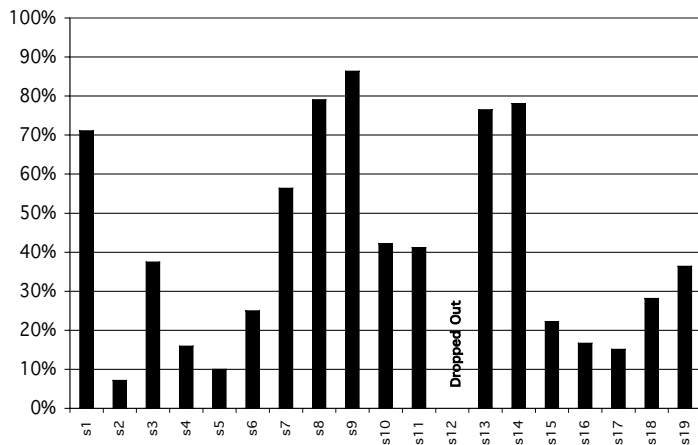


Figure 1: Study 1 Participation Rates

Nature of Reports. Participation rates, however, are not necessarily indicative of the *quality* of the responses. For example, S16, a college instructor, had a relatively low response rate of 17%. Nevertheless her reports were thoughtful, and reflective of the technology in her life:

The phone is actually becoming part of my life I guess. Kind of feeling, like, a lot more comfortable with it after that experience last week of having the phone ring during my class. I figured out about setting the ringer off, not just dialing in the ringer off, but setting it and once I got that figured out then I felt a lot more comfortable. And now I really have it with me. I made a phone call from the grocery store the other day while I was waiting for a prescription to be filled. That was kind of fun. Actually I can't remember whether I made it or got a phone call there but keeping in touch with my friend especially who has a cell phone. I really only have one friend who has got a cell phone and uses it regularly. I have a couple of others that have them but don't use them very often. And keeping in touch with him is actually- is something that I like quite a lot. Just kind of checking in since I live alone and so forth, just being able to check in with someone just at odd moments is really very comforting, I like that. And I am getting a lot more familiar with it. I still have trouble with some of the features and some of the oh- getting the, you know, getting the keypad turned off and on? I have a hard time. I am not adept yet at— I don't know, this is

interesting—at getting the call, receiving a call when the keypad is off. I know in my mind that you can just press Talk and go ahead and get the call, but somehow I disconnect myself a lot on that, so I am not sure where I am just hitting a lot of other buttons or what. I guess that is to be expected with a small keypad.

The topics participants reported on were many. Some were especially focused on usability issues while others were much more reflective about the role of mobile telephony in their lives.

Participants' diary reports sometimes hinted at issues that deserved much deeper investigation and empirical treatment. Because they were spurred by real events, the reports sometimes raised issues that did not emerge in the interviews because participants forgot about them or because we didn't know to ask about them. One investigative area that emerged in this way was development of users' mental models for long distance, roaming and digital service [17]. Here, S1 describes the use of her phone in a geographically mountainous area, where digital signal from her service provider was weak. Her report reveals deep confusion about roaming, long distance, and service-provider services—confusion that we came to learn was shared by almost all participants:

I am up in [the Colorado Rockies]. ... I called [Customer Service] and learned how to put my phone on only picking up [the service provider's] satellite so that I could call anywhere in the state for free. And then I came up to [the mountains] today and I couldn't get any service so I went back in under network and figured out how to put my phone back on analog roam. Anyway, that's the latest. I like having this option and when I go back [home], I will probably put my phone back on, just having [the service provider's] roam service.

In this next example, S3, who in the end gives an informative report, claims twice that he has “nothing to report”:

I really don't have anything new to report. I have used my phone everyday. I have about 51 or 52 numbers in my phone book now, which I find incredibly handy, really easy to use. Other than that, I don't have anything new to report, except that I still haven't been able to figure how to get voicemail messages to- for it to tell me that I have a voicemail message. It just comes up zero voice whether I've got a message or not. So that's still a mystery that I haven't solved.

Even though S3 was able to provide useful information in spite of his doubt that it would be investigatively interesting, a few participants had real difficulty doing so. Our unstructured approach meant that some participants were uncertain about what to report, even though we provided examples and assurance repeatedly. S13, for instance, had a high diary participation rate of 76%, but his reports primarily consisted of claiming he had nothing new to report, despite our repeated intervention. In the Lessons Learned section, we reflect on ways we may have improved such an outcome.

Study 2: Larger, Narrowly-Scoped Investigation

Description of the Wireless Life-Cycle Panel Study

The Wireless Life-Cycle Panel (LCP) Study was an extension of the Going Wireless study, using the findings from the Going Wireless Study to set its investigative boundaries. The LCP had a total of 200 novice wireless telephony users to be tracked throughout their first year following acquisition. A commercial telecommunications merger, however, meant that the panel could be followed for only 6 months, which still represented a serious effort. The investigation, run by the service provider's marketing and human factors divisions, employed several data collection methods, including quantitative telephone census surveys, qualitative but structured face-to-face interviews, focus groups, and a diary study. There were two primary investigators and several supporting personnel to help with recruitment logistics and survey administration. However, unlike the Going Wireless study, all investigators worked part-time on the project.

The Going Wireless study found that the first 30 days after acquisition was a time of marked learning and adjustment for new mobile phone users. To capture these rapid changes for the LCP members, we decided to again employ the voice-mail diary study method just for this initial period of time.

Application of Method: A Semi-Structured Diary

The scale of this study meant labor-intensive data collection. As such, we designed the voice-mail diary to run independently, with less investigator involvement than for the Going Wireless study. Also, the investigative issues for this study were well defined because they were informed by the results of its predecessor. Therefore, the diary study was more structured but with room for participants to elaborate, which streamlined participant reports and the subsequent data analysis. Specifically, we designed the voice-mail system to provide prompts to participants for their diary reports:

Hi, you have reached the Wireless Diary Study reporting center. Please respond to each of the following questions and then press pound.

1. Please state your first and last name.
2. Please tell us how you have used your phone service since the last time you called us, as well as any problems you have had. Also, describe any changes you may have made to your phone or service.
3. Tell us about any calls you have made to the Wireless Customer Service Center or contacts they may have made with you by email, mail, or phone.
4. To what extent is the wireless phone and service meeting your needs? Do you have any other comments or suggestions for improving the quality of your wireless service or support?

In designing the prompts, our goal was to structure responses slightly while avoiding many key presses and answer categorization. We tried to encourage breadth in the responses by minimizing the constraints that pointed questions could introduce. We were especially concerned about these issues because the diary was to span the course of a month, and ran concurrently with other data collection

procedures. In such a context, we felt that repeated interaction with a highly constrained diary would wear thin over time.

Participants were paid \$2 per call, an increase from the Going Wireless study because of the additional time required to engage with the voice-mail system. Additionally, we learned that \$1/day reward scheme garnered a reasonable response rate, and hoped that the increase would encourage additional, sustained participation among this group of participants. In this study, we did not use the "nothing to report" option, instead hoping that the prompts would encourage participants to report even when they did not use their phones. As in the Going Wireless study, participants were not explicitly told to use their mobile phones to call in, in an attempt to mitigate bias in the data.

Method Results

As with the Going Wireless study, the diary study was a useful supplement to our other data sources. Reports helped to tell a complete story: They were used to illustrate and enrich descriptions of findings that emerged from other data sources. In addition, they helped to illustrate how events and experiences unfold chronologically and to identify issues for further investigation.

Participation Rates. Overall, participation rates were lower than for the Going Wireless study. In all, 18 of the selected 20 panelists participated in the voice-mail diary (two dropped out of the LCP early on) with about a 25% participation rate across all participants (see Figure 2).

We attribute the lower participation rate to the less personal relationship investigators had with participants. In this study, participants' contact with the study personnel was limited to the initial recruitment call and the occasional phone census survey. Furthermore, the voice on the system was not that of the phone interviewer, with whom they were most familiar. We discuss both the successes and limitations of this particular instantiation of the method in the Lessons Learned section.

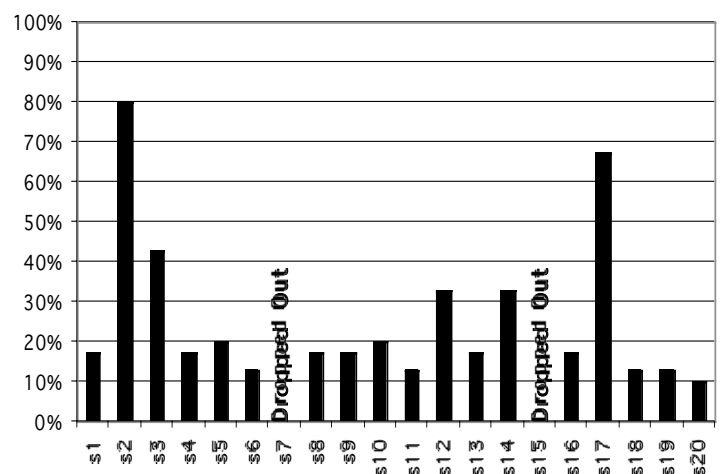


Figure 2: Study 2 Participation Rates

Nature of Reports. Although calls were less frequent than expected, they nevertheless yielded rich descriptions of activity and experiences that fleshed out findings emerging from other data sources. Reports captured participants' confusion and frustration about services and signal coverage, for example, and were far more convincing and compelling than reduced quantitative data reports alone. The diary study was also successful in soliciting design improvement ideas as they occurred to users real-time:

I would like to see...a less slippery, slim wireless phone that doesn't fall out of your hand. Can they put some sort of rubber grippers or adhesive on...? The display is also hard to read without glasses. I would like to see the letters and numbers made bigger- maybe use one color for letters, another for numbers. Lastly, I would like to have a vibrating ringer...as the ringing is a problem out in public places and meetings...

The diary reports provided highly valuable information about the rationale of behaviors and chronology of developing problems, which is hard if not impossible to capture in interviews and surveys. In this report, the participant describes her rationale for restricting mobile phone use:

I haven't received my first bill, so I haven't been using it much just because I'm not sure if I have any minutes or not.

The next series of reports, S3 shows how service coverage problems bring an initially happy customer to the brink of service disconnection over the course of less than two weeks:

Call 1: Currently the phone is meeting my needs exactly as I hoped it would. We are not having any problems. As long as the cell signal...works, I don't have any complaints.

Call 3: I contacted [the service provider] with problems. I called [customer service]-waited for 15-20 minutes. I thought the handset might be set wrong, so I had them make changes. I tried calling, re-calling, it didn't work. They called again, waited 25 minutes. The guy said [my town's] cell tower's not working ...

Problems continue to persist:

Call 7: The phone is barely meeting my needs. Only on weekends when my husband is in [another town] can he use it, so all week long when he needs to use it, he's unable because...[the service provider] has their head up their butt, and the cell towers won't work.

Call 8: The phone is not meeting my needs because it doesn't work in [the town] where we use it.

Call 9: During the week, my husband has the phone in [a town] in which it doesn't work... because [the service provider] sucks.

Happily, her problem was soon solved after this last call, but it was this type of case that provided information fundamental to our study of the early user experience.

LESSONS LEARNED

Our initial employment of the voice-mail diary method in the Going Wireless Study was itself experimental, and determining appropriate incentives was a part of the

learning experience. In deciding to reward participants \$1/day for each day they called in at least once, we were concerned about biasing the data by requiring participants to conjure up something to say even if they did not use their phones. To counter this concern, we allowed them to simply say that they had "nothing to report," which let them still earn their dollar for that day.

This was a mistake, in hindsight, and our worry about biasing the data in that particular way was needless. Although we provided the "nothing to report" option for times when participants didn't use their phones, some routinely invoked it at other times. If these participants thought that their daily phone use was mundane and not of research interest to us, they would repeatedly say that they had "nothing to report." Furthermore, it would have been useful to hear explicitly that they didn't use their phones on a particular day and why. The unexpected use of our suggested phrase, in conjunction with the uncertainty that some participants felt using the highly unstructured diary format, made inadequate diary reports from a few participants a persistent problem.

Voluntary participation with a diary study technique can be appropriate if the investigation requires multiple forms of data collection over some period of time, and there is a worry about long-term, sustained participation. It is also appropriate if the diary data is intended to flesh out the other data sources, but is not for systematic identification or quantification of issues. However, we believe that required participation will work very well, especially if the compensation schedule is tied to calling the voice-mail system a minimum number of times. Remarkably, \$1/day over a 6 week period was enough to motivate some people to consistently call in, suggesting that even as little as \$5 per day could encourage sustained participation.

The design of the voice-mail diary in the LCP study responded to the limitations of our first approach by structuring the interaction. The system prompted participants with questions, but they could respond in the negative if necessary. However, we believe that the lower participation rate—in spite of the increase to \$2/day for approximately the same duration—was due to the limited interaction participants had with one, dedicated investigator.

In summary, to encourage sustained participation, we recommend that investigators:

- Have one point of contact across the data collection activities, including the recording on the voice-mail greeting, to help develop continuity and loyalty.
- Provide periodic reminders by phone contact, especially if opportunities to meet face-to-face are limited.
- Consider creative reimbursement strategies. Make reimbursement incremental so that the total sum earned is substantial, even if the compensation per call is low. Consider issuing a base level of compensation only once a certain participation level is achieved. Providing bonuses when a certain participation level is reached might also be appropriate. However, the compensation

scheme needs to be concurrently designed with the rest of the study so that the rewards don't bias the data with unnecessary reports.

- If the experimental design permits without biasing data of research interest, program the voice-mail system phone number into participants' mobile phones and show participants how to speed dial. If hearing about activities as they occur is one of the research objectives, creating this shortcut simplifies contacting investigators.

To encourage high-quality diary reports, we recommend that investigators:

- Monitor and transcribe reports immediately to assess their quality, and provide additional training and guidance as necessary to participants.
- Provide clear examples of the range of activities that are of investigative interest. Using a semi-structured diary will prompt users each time they call in, which can be helpful but, on the other hand, can potentially constrain the elaboration of reports over time.
- Remember that talking to a voice-mail system can be intimidating for some. Provide frequent encouragement. A handout with instructions and the steps they can expect to hear from the system can help participants prepare their responses ahead of time.

Device as Subject and Means of Investigation

When studying mobile activities and technologies, employing mobile phones in the collection of voice-mail data may be especially powerful. An important issue to address is whether to ask participants to use their own mobile phones, or to distribute mobile phones. The latter option could also be an incentive for people to participate in the study.

Methodological concerns arise when the object of the study is also be the means of data collection. This was an issue for our studies. Participants were free to use their mobile phones to call in while in new and different situations, but, because we were studying usage patterns, we did not want to overtly influence this decision so we never suggested they use their mobile phones. Brown, O'Hara, and Sellen [2] had similar concerns in their study of information capture. In distributing cameras to their participants to visually record their own information capture activities, they instructed that the cameras were to record the activity and not be the means for information capture.

In some cases, it could be important for a research study to *encourage* use of the mobile phones, even if they are under study. For example, usability specialists may distribute mobile phones for evaluation, and deliberately instruct participants to call the voice-mail diary with their mobile phones to report problems as they occur.

IMPLEMENTATION CONCERNS

Setting-Up the Voice-Mail System

In setting up a voice-mail diary, it is important to consider system constraints before designing the diary. Setting up a voice-mail diary will require working with voice-mail system administrators. Before commencing diary design, it

is important to investigate the degree of flexibility the voice-mail system and administrators allow. Issues to consider include:

- *Mailbox Type.* Voice-mail systems can have different types of mailboxes. Some are simple mailboxes that capture speech, and others are designed to provide prompts and receive keypad input. Determine that the voice-mail system can support these different types before designing the diary procedure.
- *Message Length.* System message length defaults may not long enough for diary reports. Ensure that the recording time is adequate. Future participation could be reduced if participants are cut off by the system.
- *Mailbox Memory.* Messages cannot be recorded when the mailbox reaches memory limits. Ascertain that memory size is large enough, especially when studies span long holiday weekends.
- *Message Life.* Some systems automatically expire messages after a certain length of time. Know and/or adjust those settings in advance.
- *Data Protection.* Voice-mail systems are not archival. Mailboxes are not usually backed up; once a message is deleted, it may be irretrievable. Consider copying messages to tape to protect the data, and to use for replaying to relevant technology developers or for other presentation purposes. High-quality transcriptions also protect data.

Transcription

Many voice-mail systems have built-in features that are useful for transcription, including play and pause, restarting, 10 second "rewind" and advance, and replay speed control. It is usually possible to retrieve the exact day and time of the recording as well.

Having had experience with video and audio transcription, we were surprised at how easy it was to transcribe the voice-mail reports. We approximate the time required for accurate transcription (although not one including pause-level timing) to be about 5 times longer than the recording itself. When transcription is done regularly, the task stays manageable.

Other Uses of the Technology

Voice-mail systems can make other logistical aspects of the research easier as well. The Panel Study showed how voice-mail menus can be used for the selection and categorization of responses. Additionally, the greeting the participants hear when dialing into the voice-mail system can be used to provide instructions for the voice-mail reports, but it can also be used for other instructions or reminders for the project as a whole. For example, we reminded participants in the Going Wireless study to contact us as soon as they received their first phone bill, which arrived on different dates across participants. It was much more convenient and efficient to do it this way for both participants and investigators.

Another suggestion is to use caller ID to track origination of call, when such technology permits. This supplementary data could be useful, depending on research objectives.

When Voice-Mail Is Not an Option

We note that some voice-mail systems or organizations using voice-mail systems provide little flexibility in configuring the system to investigators' needs. Although using voice-mail as a recording medium provides many benefits by allowing the investigator to handle multiple simultaneous incoming calls, easily change greeting messages, easily access diary entries from anywhere, and transcribe messages using built-in replay features, other audio recording media can be used when voice-mail system opportunities are limited. Answering machines can be used, although callers might get a busy signal if another person is calling at the same time on the same line (multiple lines and answering machines might be another solution, but one imagines that if voice-mail configuration is difficult, installing multiple phone lines probably is, too!). However, some answering machines have automatic time-stamping of messages and allow the investigator to dial in remotely. Individual audio recorders can also be used, but they introduce some of the same limitations of paper diaries that we tried to avoid: they require participants to carry an additional device dedicated only to data collection. Also as with paper diaries, audio recorders require that participants and investigators rendezvous more frequently to collect data, and automatic time-stamping is not possible. Still, when voice-mail is not an option, we encourage investigators to consider how these other recording technologies might be adapted to collect audio diary entries, which can be richly compelling.

VARIATIONS ON THE METHOD

Structured v. Unstructured Overview

This paper has reviewed unstructured and semi-structured versions of the voice-mail diary study method. Our unstructured use of the method was almost stream-of-consciousness in style. The narratives collected were very useful, especially for shaping high-level investigative objectives. Some participants were uncomfortable with the unstructured style, however, and their reports were only minimally useful. The second study addressed this problem by using a semi-structured diary, providing prompts to guide users to talk about particular issues, although still allowing for elaboration.

A highly structured diary would afford systematic collection of quantitative data collection over a constrained issue set. A highly structured diary could be designed as a set of menu options that accepts keypad selection input. Thoughtful attention would be required, however, in making such an interaction a pleasant one for participants. The more structured a diary, the more testing and design iteration would be required to determine appropriate selections. Therefore, we recommend that a highly structured diary be used only after an initial study clearly defines the relevant investigative issues, and that the same guidelines for survey preparation be followed for diary design. Finally, structured diaries should include at least

one open-ended question to allow participants to qualify answers or address issues beyond those specified in the diary.

Semi- or highly- structured diaries are useful for larger study populations, for collection of more quantitative versus qualitative data, and for when time for participant contact, transcription and analysis is limited. In designing the diary, attention should also be paid to the amount of work required for participation: Diary participation requirements should be balanced with those of other data collection activities.

Participant- v. Investigator- Directed Approaches

Our use of the diary method required participants to remember to make a diary entry. We call this a participant-directed approach. However, mobile telephony affords innovative, investigator-directed variations on diary methods that are inspired by ESM [10], described earlier. Using voice calls, text messaging, and even handset calendars, investigators can remind or even decide when participants record activity. Like the pagers used in ESM, text messages can be sent to participants' mobile phones to signal that they are to call in to the voice-mail line. The text messages can be sent to ask specific questions that participants could respond to in the voice-mail report. Similarly, investigators can call participants at times of interest to the investigators to conduct a short "impromptu" phone interview for which responses would be highly situated. Finally, calendars on phones could be programmed to automatically alert and remind participants when to call in to the voice-mail diary.

CONCLUSION

Multiple research objectives and analytical approaches can be satisfied when using the voice-mail diary technique. When paired with mobile telephony, opportunities for creative experimental design only increase, and make the technique especially suitable for studying activity naturalistically and with minimal intrusion under mobile conditions. Voice-mail diaries are a low-cost way to collect data about activities as they occur at any time, with other people, and in any number of settings, including ones where investigators would not normally or easily be permitted. Participant- and investigator-directed approaches allow additional flexibility in the manner in which that data is captured.

ACKNOWLEDGMENTS

We thank Ed Youngs, Carolanne Fisher, Julie Kopp, and Larry Durkin for their support of our research on mobile telephony. Abi Sellen, Don Norman, and Jurek Kirakowski provided useful information about the history of the diary study method. Beki Grinter, Jonathan Grudin, Ellen Isaacs, Tom Moran and Abi Sellen and provided helpful reviews of earlier versions of this paper. This work was supported in part by a grant from the National Science Foundation #9977952.

REFERENCES

1. Adler, A., Gujar, A., Harrison, B., O'Hara, K. and Sellen, A. (1998). "A Diary Study of Work-Related Reading: Design Implications for Digital Reading

- Devices," *Proceedings of ACM CHI 98 Conference on Human Factors in Computing Systems*, v.1 p.241-248.
2. Brown, B., Sellen, A. & O'Hara, K. (2000). "A Diary Study of Information Capture in Working Life," *Proceedings of ACM CHI 2000 Conference on Human Factors in Computing Systems*, v.1, p.438-445.
 3. Carayon, P. and Hajnal, C. (1993). "A Diary Study of Computer Use and Worker Stress: Preliminary Results VI. Longitudinal Studies," *Proceedings of the Fifth International Conference on Human-Computer Interaction*, v.1 p. 715-720.
 4. Chin, J. P., Herring, R. D. and Familant, M. E. (1992) "A Usability and Diary Study Assessing the Effectiveness of Call Acceptance Lists," *Proceedings of the Human Factors Society 36th Annual Meeting*, v.1 p. 216-220
 5. Csikszentmihalyi, M., Larson, R., & Prescott, S. (1977). The ecology of adolescent activity and experience. *Journal of Youth and Adolescence*, v. 6, p. 281-94.
 6. Csikszentmihalyi, M., Rathunde, K. and Whalen, S. (1993). *Talented Teenagers: The Roots of Success and Failure*. New York: Cambridge University Press.
 7. Eldridge, M. and Newman, W. (1996) "Agenda Benders: Modelling the Disruptions Caused by Technology Failures in the Workplace," *Proceedings of ACM CHI 96 Conference on Human Factors in Computing Systems*, v.2 p.219-220.
 8. Rebecca E. Grinter and Margery A. Eldridge (2001). "y do tngs luv 2 txt msg?" *Proceedings of 7th European Conference on Computer Supported Cooperative Work (ECSCW '01)*, Bonn, Germany, pp. 219-238.
 9. Kirakowski, J., and Corbett, M. (1990). *Effective Methodology for the Study of HCI*. New York: Elsevier.
 10. Larson, R. & Csikszentmihalyi, M. (1983). The Experience Sampling Method. In H. T. Reis (Ed.), *Naturalistic approaches to studying social interaction: New Directions for Methodology of Social and Behavioral Science*. San Francisco, CA: Jossey-Bass.
 11. Mark, G. and Christiansen, U., and Shafae, M. (2001) "A Methodology Using a Microcamera for Studying Mobile IT Usage and Person Mobility," Presented at the CHI Workshop on Mobile Communications: Understanding Users, Adoption & Design, *ACM CHI 01 Conf. on Human Factors in Computing Systems* (see http://www.cs.colorado.edu/~palen/chi_workshop/papers/Mark.pdf)
 12. Newman, W., Eldridge, M. and Harper, R. (1996). "Modelling Last-Minute Authoring: Does Technology Add Value or Encourage Tinkering?," *Proceedings of ACM CHI 96 Conference on Human Factors in Computing Systems*, v. 2 p.221-222
 13. Norman, D. A. (1981). Categorization of action slips. *Psychology Review*, v. 88 (1), 1-15.
 14. O'Hara, K., Perry, M., Sellen, A. and Brown, B. (to appear). "Managing Information on the Move: A Study of the Information Needs of Mobile Professionals," In B. Brown, N. Green, and R. Harper (Eds.), *Wireless World: Social and Interactional Aspects of the Mobile Age*. Berlin: Springer-Verlag.
 15. O'Hara, K., and Perry, M. (2001). "Shopping Anytime Anywhere." *Extended Abstracts Proceedings of the 2001 ACM Conference on Human Factors in Computing Systems (CHI '01)*, Seattle, WA, p. 345-346.
 16. Palen, L., Salzman, M. & Youngs, E. (2000). "Going Wireless: Behavior and Practice of New Mobile Phone Users." *Proceedings of the ACM 2000 CSCW 2000 Conference on Computer Supported Cooperative Work*, Philadelphia, PA, p. X.
 17. Palen, L. and Salzman, M. (in press) "Beyond the Handset: Designing for Wireless Communications Usability," *Transactions on Computer-Human Interaction*.
 18. Rieman, J. (1993). "The Diary Study: A Workplace-Oriented Research Tool to Guide Laboratory Efforts Collecting User-Information for System Design." *Proceedings of the ACM INTERCHI'93 Conference on Human Factors in Computing Systems*, p.321-326.
 19. Robinson, R. (1996). "Panel The Participatory Design of Work Space," Harrison, S. (Panel Organizer), *Proceedings of the 1996 Conference on Participatory Design*, Cambridge, MA.
 20. Sellen, A.J. (1994). Detection of everyday errors. *Applied Psychology: An International Review (Special issue on error detection)*, v. 43(4), 475-498.
 21. Sellen, A. & Harper, R. (1997). "Paper as an Analytic Resource for the Design of New Technologies," *Proceedings of ACM CHI 97 Conference on Human Factors in Computing Systems*, v.1 p.319-326.
 22. *The Sonic Memorial Project: A Nationwide Public Radio Collaboration* (Launched October 2001), <http://www.sonicmemorial.org/>.