Using the Experience Sampling Method to Evaluate Ubicomp Applications

Ubiquitous computing's overarching goal is for technology to disappear into the background yet remain useful to users. A technique from the field of psychology—the Experience Sampling Method—could help researchers improve ubiquitous computing's evaluation process.

Sunny Consolvo and Miriam Walker
Intel Research Seattle

The evaluation process is critical for successfully deploying new technologies. You can use evaluations throughout a technology's development to inspire new applications by identifying unmet user needs and to help learn whether your applications meet the needs of your users. If your applications are not meeting your users' needs, evaluations can help identify the reasons.

There is a class of ubicomp devices designed to accompany users through different tasks and in changing contexts and environments. Designing for this class is particularly challenging, especially given the variability of users, uses, and environments involved. For example, a mobile phone user is likely to use the device at the office, at lunch, in the car, at the store, and at home. Such settings cannot be reasonably approximated in a traditional laboratory.

Recently, the Intel Research Personal Server team wanted help creating realistic usage scenarios. The Personal Server is a mobile device that lets the user store and access data and applications through interfaces found in the environment. The Personal Server does not use a display; it wirelessly connects to nearby input and output devices.[1] To help create the usage scenarios, the Personal Server team wanted to gain a better understanding of people's information needs. Our challenge as evaluators was to find a technique that would help the Personal Server team collect appropriate data. Given that the Personal Server application has a broad group of target users, we needed an evaluation technique that would accommodate many participants in several environments.

People will adopt and use ubicomp applications—including the Personal Server—in several settings and for potentially different tasks, so appropriate evaluation techniques must take place in those settings and explore those different tasks. Such requirements are often called *in situ* (in the actual situation) or *ecologically valid* ("the occurrence and distribution of stimulus variables in the natural or customary habitat of an individual").[2] The evaluation techniques used for these ubicomp applications would ideally take place in situ, involve several participants, take place over time, and collect both qualitative and quantitative data. To help with the Personal Server evaluation, we used a technique from the field of psychology—called the Experience Sampling Method[3]—that other researchers have found to be effective for learning about situations and person-situation interactions. The technique compares most closely with recall-based self-reporting techniques such as interviews, traditional surveys, and diaries.

Experience Sampling Method

In other contexts, researchers have called ESM by various names, including *time sampling*, *beeper*

Published by the IEEE CS and IEEE ComSoc ■ 1536-1268/03/$17.00 © 2003 IEEE

TABLE 1
Major alerting choices for the researcher.

Categories	Considerations	Description	Drawbacks
Type of alert	Random	Alerts are delivered randomly.	Requires a flexible tool.
	Scheduled	Alerts are delivered on specified schedule.	Might introduce cognitive bias.
	Event-based	Alerts are triggered when event of interest occurs.	Might introduce cognitive bias if triggered by participant.
Scheduling requirements	Daily time period	Deliver alerts within a specified daily time frame only.	Requires a flexible tool. If time frame isn't well chosen, might miss interesting situations.
	Number of alerts per day	Deliver a specified number of alerts per day	Requires a flexible tool. Might introduce bias if participants know how many alerts they will receive.
	Number of alerts overall	Deliver a specified number of alerts for the study's duration.	Requires a flexible tool.
Delivery mechanism	Audible	Deliver alerts the participant can hear.	Might be inappropriate in certain situations (such as a movie theater or meeting).
	Tactile	Deliver alerts the participant can feel.	Participant might not feel alert. Requires a flexible tool.

study,[4] and *ecological momentary assessment*.[5] Essentially, participants fill out several brief questionnaires every day by responding to alerts. ESM does not require participants to recall anything; instead, questionnaires ask about the participants' current activities and feelings. Although ESM is a self-report technique, the no-recall feature reduces the cognitive biases associated with other recall-based self-report techniques such as interviews, traditional surveys, and diaries. Researchers have traditionally used ESM to understand areas such as mood, time use, and social interactions.

The questionnaires sample participants' experiences throughout the day. Researchers are not present during the ESM evaluation. Researcher involvement typically takes the form of interviews at the beginning and end of the study. This minimal involvement reduces biases associated with participants being observed. It also reduces cost and facilitates involving several participants. ESM studies often involve 30 to 80 participants over one to three weeks, with each participant receiving up to 10 alerts per day. A nice feature of ESM is that you can use statistical methods to evaluate the data, a process that is often not possible with studies involving small numbers of participants. In addition, you can collect structured data (fixed responses used to generate quantitative data) and unstructured data (open-ended questions used to generate qualitative data). Because data can be collected over a long period of time, researchers might be able to capture infrequently occurring events. Because no researcher is present, it might be possible to learn about events not normally accessible to an observer (for example, getting ready for work in the morning or going on a date).

You have to make several choices prior to beginning an ESM evaluation. These choices depend on the study's individual needs. For example, you must choose how to alert participants that it's time to complete a questionnaire. You must also choose how to deliver the questionnaires and capture participant responses. When making these choices, you must consider the burden on the participants and the constraints of their environment. For example, asking many open-ended questions several times per day would be cumbersome. A good rule of thumb is to ensure that participants can complete the questionnaires in less than two minutes. One way to control the burden of open-ended questions on participants is to use probabilities in your question design. However, if you must ask several open-ended questions every time, you might consider alerting the participant fewer times per day or using a technique more appropriate to open-ended questions.[6]

Alerting

When choosing how to alert the participant, you must choose the type of alert, the scheduling requirements, and the delivery mechanism. The types of alerts (listed in Table 1) include *random* (for example, 10 randomly spaced alerts every day), *scheduled* (for example, 10 evenly timed alerts every day), and *event-based* (when a particular event of interest occurs, alert the participant). You cannot rule out bias for scheduled and participant-triggered event-based alerts. With scheduled alerts, participants can anticipate when an alert will arrive and might modify their behavior accordingly; with participant-triggered event-based alerts, you have no guarantee that the participant will trigger the questionnaire when the event actually occurs.

When it comes to scheduling requirements, you must consider whether you need to specify the time period during which alerts should be delivered. For example, if there is no compelling reason to wake the participant in the middle of the night to complete a questionnaire, you might want to deliver alerts only within a certain daily timeframe. If participants' schedules will be different on weekdays than on weekends, you might want a tool that offers the flexibility to specify different schedules for different days or for different participants. You might also want to control how many alerts you deliver each day. For example, you might want to specify that 10 alerts should be delivered per day, with a study total of 140 alerts. Specifying a study total means that even if the participant cannot return the

TABLE 2
Data delivery options for the researcher.

Categories	Considerations	Description	Downsides
How to deliver	Written questionnaires	The participant reads the questionnaire from a screen or paper.	Font size, contrast, and available lighting could impact the participant's ability to read the questionnaires.
	Audible questionnaires	The participant listens to the questionnaire.	Participant might not be in an appropriate place to listen to the questionnaire out loud. If using a phone, service availability might be a problem. Listening to questions will likely take longer than reading them. Multiple choice answers require the participant to remember the choices.
Question design	Question order	Should the questions be in fixed or random order?	If random, requires flexible tool.
	Probabilities	Questions might be assigned a probability that they will be asked.	Requires flexible tool.
	Contingencies or branching	Should questions be asked based on certain responses or on other questions that have been asked?	Requires flexible tool.

equipment immediately upon completion of the study, the participant will not continue to be alerted.

You must also choose how to deliver the alerts. To make this choice, you should consider the participants' environments. For example, if the participant is likely to be in a theater or in meetings, an audible alert might be inappropriate. If the participant is likely to keep the alerting device tucked away in a bag, a tactile alert might be of no use. If you can't predict where the participant is likely to be, you might need a tool that lets the participant change the type of alert at any time during the study. Tools commonly used to alert the participant include mobile phones, pagers, watches, PDAs, and custom devices.

Delivering

When choosing how to deliver the questionnaires, you must consider several options, many of which are listed in Table 2. For example, should the participant listen to or read the questionnaires? If the participant reads the questionnaires, are there font size or contrast issues? If the participant listens to the questionnaires, will environment noise be a problem? If the participant plays the questionnaires aloud, will the sound disrupt people nearby? Other considerations concern question design. For example, should the question order be fixed or random? Do you need to use contingencies based on questions asked or responses given? Do you need to assign probabilities that questions be asked?

Tools commonly used to deliver questionnaires to the participant include phones (both traditional and mobile), paper booklets, PDAs, audio recordings, and custom devices. For evaluations that use phones, the participant would call a designated phone number upon receiving an alert. For evaluations that use an audio recording, a participant might listen to a prerecorded questionnaire on a device such as a cassette player.

Capturing

When choosing how to capture participant responses (see Table 3), you must choose whether to collect structured or unstructured data—a decision that might affect whether the participant should speak or write the responses. For example, audio recordings are particularly appropriate for answering open-ended questions but they require the participant to speak, which is less convenient than writing in many circumstances. In Leysia Palen and Marilyn Salzman's voicemail diary studies, participants calling a recording service left long and detailed accounts.[6] Spoken responses were not an issue in that study because participants called in at their convenience rather than being alerted randomly. Writing has drawbacks too. In addition to limiting how much a participant might say to an open-ended question, writing is difficult to do while driving or walking.

You must also make choices concerning timing and data transfer by considering, for example, whether it is important to know how soon the participant responds after receiving an alert. You might also want the ability to put a limit on how much time a participant has to respond to an alert before considering the alert missed. And you might want to limit the amount of time a participant has to respond to individual questions before considering the questionnaire incomplete. Paper-based tools handle such requirements poorly. While PDAs handle such requirements well, they might intimidate some participants.

The speak-or-write decision will affect data transfer. In most cases, you will want the data in an electronic format so you can use a data-analysis tool. Audio recordings require transcription, a process that can introduce errors. Handwritten responses from paper-based tools also require transcription, which can introduce errors. Automatic voice- or handwriting-recognition approaches are not mature enough to use

TABLE 3
Data capturing options for the researcher.

Categories	Considerations	Description	Downsides
How to record the responses	Written responses	The participants write their responses to the questions.	When responding to open-ended questions, participant responses might be short. In some circumstances, writing responses may be difficult.
	Spoken responses	The participants speak their responses to the questions.	Participants might not be in an appropriate place to speak their responses out loud. The responses have to be transcribed to electronic format, which might introduce error.
Timing of responses	Response time	Time it takes participant to respond to the alert.	Requires flexible tool.
	Time-out	Mark questionnaire "missed" or "incomplete" if participant doesn't respond to the alert or an individual question within a specified amount of time.	Requires flexible tool.

because they have unacceptably high error rates. However, PDAs store responses in a simple electronic text format, eliminating transcription error. Commonly used tools for capturing participant responses include paper booklets, PDAs, phones (traditional or mobile), audio recordings, cameras, and custom devices. In the case of phones, the participant calls a designated phone number, then speaks the responses to a live researcher or to a recording service. In the case of audio recordings, the participants could speak their responses into a recording device, such as a cassette player.

In general, the type of tools you use will depend on your study's needs. Table 4 summarizes commonly used tools that range in cost and flexibility. Using fewer tools should mean less burden on the participants. The more flexible tools often introduce issues of cost, power consumption, and sometimes service availability. And with expensive tools, you run the risk of them being damaged, lost, or stolen.

PDA-based ESM

Our main experiences with ESM involved PDAs running the PalmOS. Based on our and other researchers' experiences,[5] we discovered some advantages and limitations of PDA-based ESM. By using PDAs, we did not have to transcribe any data. The PDAs transfer all data to the analysis computer through the PDA's synchronization facilities. PDAs also give us precise control over alerts, the ability to specify the daily time period during which alerts should be delivered, the number of alerts we can deliver per day, and the time window within which participants must respond to alerts. Depending on the device's features, we can deliver alerts with an audible beep, a flashing LED, a tactile vibration, or some combination of the three.

We can also choose random, scheduled, or participant-triggered event-based alerts. The software for running ESM on PDAs can capture when alerts are delivered and when, if at all, participants respond within the specified window. PDA-based ESM also offers flexibility with question design, so we can include both structured and unstructured questions in the same study. Finally, we can deliver questions in fixed or random order and take advantage of probabilities and contingencies. However, despite these strengths, PDAs might be a burden for some participants. If the participants have never used a PDA, it might be intimidating. They must also carry the device at all times during the study. Some participants, particularly those with vision or motor-skill impairments, might have difficulty using PDAs.

Another issue is that PDAs do not offer stable data storage. If the batteries run out or fall out, the data disappears. Relying on participants to charge the batteries or upload the data is risky. For battery conservation purposes, we used PDA software that locked the participant out of all PalmOS features except the ESM study. Finally, we risk expensive equipment being damaged, lost, or stolen. Despite these limitations, we chose PalmOS PDAs because of the flexibility they offer to evaluators and participants. We were nervous about having to rely on participants to speak their responses, and we also did not want the added expense and availability problems associated with mobile phone or pager services.

TABLE 4
Commonly used ESM tools.

	Alerting	Delivering	Capturing
PDAs	x	x	x
Paper booklets		x	x
Mobile phones	x	x	x
Traditional phones		x	x
Audio player/recorder		x	x
Pagers	x		
Watches	x		
Cameras			x
Custom devices	x	x	x

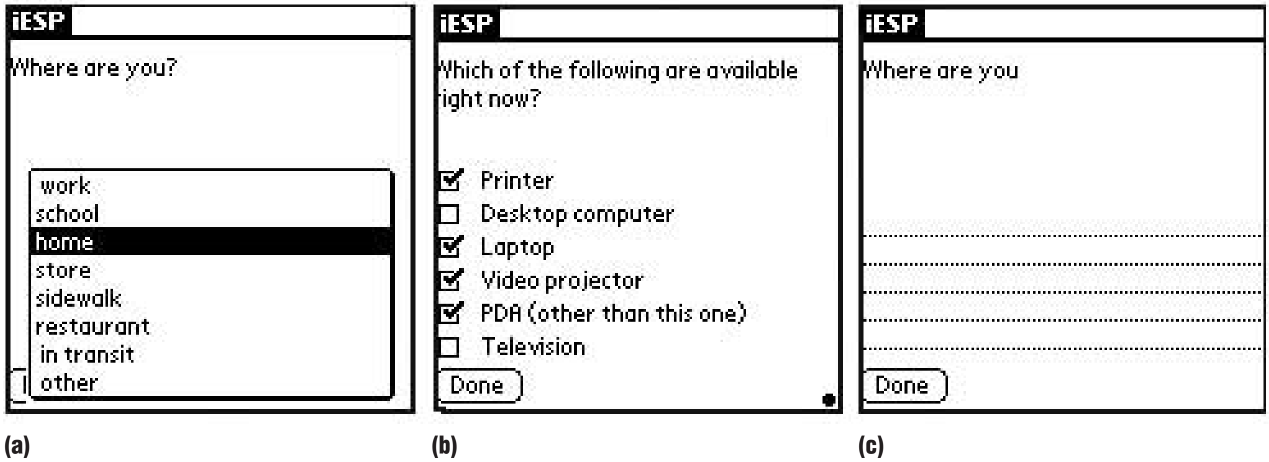

Figure 1. (a) Example of a multiple-choice, single-answer question. (b) Example of a multiple-choice, multiple-answer question. (c) Example of an open-ended question.

Our experiences

For the Personal Server's Information Use Study, we collected data including participants' information needs and what output devices were available to them throughout the day. We conducted the study in August and September 2002. The study involved 31 participants (21 female, 10 male) from the local Seattle community. We recruited participants by hanging posters in shops and restaurants, posting announcements on an Internet message board, and word of mouth. We required that participants be over the age of 18—their actual ages ranged from 18 to 75—and that they use a mobile phone regularly. Our participant group represented a variety of occupations including administrative assistants, hotel employees, real estate agents, sales and marketing professionals, teachers, and college students.

The study involved hour-long pre- and post-study interviews and seven days of ESM testing in the field. For the ESM phase, participants received 10 randomly scheduled alerts per day in a 12-hour time window. For 30 of the 31 participants, the daily time window was 9:00 AM to 9:00 PM. For the other participant, the daily time window was 12:00 PM to 12:00 AM because that participant worked a night shift. Each time window consisted of 72-minute intervals; one alert occurred randomly within each interval. We offered the participants an incentive of $50 for participating and returning the equipment plus $1 for each completed questionnaire. The total possible incentive was $120. We were hoping that the per-questionnaire incentive would encourage participants to complete more questionnaires. We paid the incentive at the end of the post-study interview after the equipment had been returned.

We used Palm m500s running iESP to alert participants, deliver questionnaires, and capture responses. Figure 1 shows examples of how the questions appeared to participants. The iESP software (http://seattleweb.intel-research.net/projects/ESM/iESP.html)—a free, open-source package for running ESM on PalmOS PDAs—is our modified version of ESP (www2.bc.edu/~barretli/esp). We also used Intel Pocket PC cameras, because each questionnaire requested that participants take a photo. During the pre-study interviews, we told participants that we understood there would be situations in which taking a photo was inappropriate. For open-ended questions, all participants could use a paper booklet, the Palm's Graffiti style of text entry, or the Palm's soft keyboard (only one participant used the paper booklet). We gave nine participants a full-sized foldout keyboard for the Palm. We chose the Palm m500 rather than a cheaper model because we encountered problems during a pilot study with the Palm m100. The m100 does not offer a vibrate mode and has a fragile battery storage mechanism—both of which resulted in serious problems.

All 31 participants returned the equipment, although one participant returned the Palm after performing a hard reset that erased his data and the iESP software. We believe this participant performed the hard reset out of curiosity and not because of malicious intent or an unwillingness to share the data. Although he took more photos than the other participants, we could not include his data in our results. One Palm was returned with a broken screen, which was the result of being dropped. All other equipment was returned in good condition. Of the 70 questionnaires per participant, the average number of completed questionnaires was 56, with a median of 58. The range was 20 to 68; 28 participants answered at least 50 questionnaires. The average number of photos taken per participant was 52, with a median of 56 and a range of 2 to 68.

During the post-study interviews, par-

Figure 2. Photos taken during the study included (a) pets, (b) vacation spots, (c) dinner preparation, and (d) the grocery store.

ticipants said that their primary reasons for not completing questionnaires included being in an inappropriate situation or not noticing the alert. Participants could control the beep volume and type of alert at any time during the study, but they complained that they often missed alerts when they were in noisy places like restaurants and festivals, despite using the beep-and-vibrate alert with the beep set to the highest volume. Inappropriate situations included talking on the phone, giving formal presentations, serving customers, and being in the middle of a job interview or a meeting with a client. Participants often focused their photos on friends, coworkers, pets, picturesque scenes, the participants themselves, their homes, shops, and so forth. Figure 2 shows a few examples.

Of the 31 participants, 28 said they would participate again, two would only do it for a higher incentive, and one professional who dealt with clients throughout the day would not do it again under any circumstances. The biggest complaint about the camera was that participants were concerned about violating other people's privacy. Other complaints included difficulty using the camera (despite a brief training session during the pre-study interview), it was an additional thing to carry, and participants thought their photos would be boring. Despite the participants' concerns, we found the photos to be very useful. They helped paint a nice picture of the participants' lives and what was important to the participants—something not as easily done from the data alone. Regarding the choice of the Palm, most participants appreciated its discreteness. As we suspected, most participants claimed that they would have completed fewer questionnaires if they had been asked to speak the responses, largely because they were often alerted in situations where speaking would have been problematic. Only one of the 31 participants would have preferred to speak the responses.

Our results

Almost 90 percent of the time, participants were either alone or talking with one other person. When they wanted to communicate with someone who was not present, the participants were usually at home, at work, in transit, or at someone else's home. When they were at locations like a store or a restaurant, the participants typically only wanted to communicate with someone who wasn't present a total of 11.6 percent of the time. This means that the Personal Server scenarios should involve situations in which users are alone or with one other person. It also means that scenarios involving communication with someone not physically with the user should be situated at home, at work, at someone else's home, or while the user is in transit.

Devices such as the Personal Server depend on a future where users have access to several output devices (shared or personal). We found that participants had access to at least one output device about 70 percent of the time, including televisions, desktop computers, printers, laptop computers, PDAs, and video projectors. The question included a reminder that the PDA being used for the study should not be considered an available PDA. It is reasonable for the Personal Server team to create scenarios where the user takes advantage of an available output device, particularly if the device is a television, desktop computer, or printer (the most frequently available devices by far). Interestingly, the video projector, a device often used in scenarios of ubiquitous computing devices, was only available to the participants 2.12 percent of the time.

We did use open-ended questions in our study, although we did not include them in every questionnaire. For example, about half of the time, we asked participants if there was any information about their current conversation, location, or activities that they wanted to know but didn't have available. Just over 7 percent of the time, participants wanted several types of information. We grouped their responses into several categories:

- Arts and entertainment
- Health, nutrition, and fitness
- Local events
- Personal information management

Related Work

We are not the first researchers to explore ESM for human-computer interaction. Stephen Intille and his colleagues explored image-based experience sampling and reflection, a technique that automatically triggers event-based alerts.[1] In that study, events of interest occurred in a single, instrumented room. Their technique does not directly relate to a study like ours, which tracks several participants throughout their day.

In another study, James Hudson and his colleagues conducted ESM with RIM Blackberry devices to explore attitudes about availability and interruptibility of managers at IBM Research.[2] Scott Hudson also used ESM to gauge the interruptibility of persons with high-level staff positions at Carnegie Mellon University.[3]

Although not ESM, Palen and Salzman's work with voicemail diaries[4] explores similar issues with evaluating mobile devices and offers examples of the types of open-ended responses that can result from using a mobile phone as a capture tool. For an anecdotal comparison of the difference in responses between answering an open-ended question with a mobile phone versus a Palm, compare Palen's excerpts to the free-form participant responses we listed in the "Our results" section.

REFERENCES

1. S. Intille, C. Kukla, and X. Ma, "Eliciting User Preferences Using Image-Based Experience Sampling and Reflection," *Proc. Conf. Human Factors and Computing Systems: Extended Abstracts*, ACM Press, 2002, pp. 738–739.
2. J.M. Hudson et al., "'I'd Be Overwhelmed, But It's Just One More Thing To Do': Availability and Interruption in Research Management," *Proc. Conf. Human Factors and Computing Systems*, ACM Press, 2002, pp. 97–104.
3. S.E. Hudson et al., "Predicting Human Interruptibility with Sensors: A Wizard of Oz Feasibility Study," *Proc. Conf. Human Factors and Computing Systems*, ACM Press, vol. 5, no. 1, April 2003, pp. 257–264.
4. L. Palen and M. Salzman, "Voice-Mail Diary Studies for Naturalistic Data Capture under Mobile Conditions," *Proc. Conf. Computer Supported Cooperative Work*, ACM Press, 2002, pp. 87–95.

- Science and technology
- Shopping and classifieds
- Transportation and locations
- Travel and outdoors
- Weather
- Unanswerable questions
- Irrelevant answer

The "irrelevant answer" category was responsible for 33 percent of responses. These tended to be responses explaining what the participant was doing, a comment about the study equipment, or a random response such as "ARGGGGGGHHH-HHH." "Transportation and locations" and "shopping and classifieds" were the second- and third-most requested. Weather, travel, and local events tied for least-requested categories. Examples of unedited responses included:

- "we want to know where happy hour bar is and where we can get waffle ice-cream cones"
- "where is this stupid restaurant"
- "I am stuck in traffic & want to know the best route home"
- "bus info"
- "Ferry schedules would help to plan a weekend visit"
- "in store shopping specials"
- "med records"

About half of the time, we asked participants if there was any information from their current location or activity they would like to save for future use. We received a wide range of responses. Some examples of things participants wanted to save that might make realistic scenarios for the Personal Server include the following unedited responses:

- "What night does this restaurant have their paneer specials"
- "i'm training on the job today, i'd like to be able to save some notes on it."
- "time i spent running"
- "i would like to save copy of this song so we can remember it and hear it later."
- "a reminder to buy the CD that I was just discussing"
- "remember location of grocery store"
- "why is my cellphone roaming"
- "photo of lake union"

This is just a small sample of the data we collected, but it shows how the combination of quantitative and qualitative data can help the Personal Server team create realistic usage scenarios. The fact that these usage scenarios will be based on real data might help convince management of the project's viability. It also increases the chance of the device being adopted by real users.

Given some recent advances in technology, we think ESM has the potential to make a real impact on ubicomp. (See the "Related Work" sidebar for leads to other research projects in this area.) Lisa and Daniel Barrett[5] have noted how wireless networking might change PDA-based ESM. For example, with wireless networking, participants could transfer their responses immediately to the evaluator, reducing the chance of data loss and letting the evaluator intervene if a participant is having difficulty with the study. Questionnaires could also be served from a single machine, removing the need for evaluators to download the questionnaires to every PDA. However, wireless networking means increased cost. It also introduces service availability problems. We suspect that being able to transfer responses immediately to evaluators might also create feelings that the participant is being watched, potentially introducing observer bias.

In addition to believing that ESM can have a measurable impact on ubicomp, we believe that ubiquitous computing applications will likely enhance ESM. For example, researchers could build ubicomp applications to identify events of interest. Such applications might remove the bias currently associated with event-based alerts triggered by participants. Ubicomp applications could also automatically collect some of the data currently asked of participants. For example, a ubicomp application could automatically capture the participant's location.

New devices are arriving on the market that could be used for ESM. For example, devices combining PDAs and digital cameras could eliminate several of the problems we experienced with cameras. With the new devices, participants would only have to carry one device, and the photos could be more easily synchronized with the appropriate questionnaire.

ESM is an ecologically valid user study technique that provides the opportunity for collecting quantitative and qualitative data. We see ESM being used as a formative technique to reveal where ubicomp solutions might be appropriate and to help define requirements for ubicomp solutions. We also see it being used as a summative technique to measure the affect of ubicomp applications on people's lives. Given that ubicomp devices like the PDA are helping make ESM more compelling, we also see ESM inspiring ubicomp applications that could help to improve the technique, making it even more appropriate for studying ubiquitous computing applications. P

ACKNOWLEDGMENTS

We thank Micah Brodsky of the University of Washington for coding iESP, Lisa Feldman Barrett and Daniel Barrett for their help with ESP, and Ken Anderson of Intel People & Practices Research for his advice on conducting ESM. We also thank Justin Silves, Gaetano Borriello, Ken Smith, and the rest of our colleagues at Intel Research Seattle for their advice, assistance, and patience while we explored this technique. Thanks also to Rahul Shah for making the data analysis easier, our pilot testers, the Intel Research Personal Server team for their inspiration (in particular, Roy Want, Trevor Pering, and John Light), and James Landay, Ken Fishkin, and Steve Gribble for their advice.

the AUTHORS

Sunny Consolvo is a member of the research staff at Intel Research Seattle. Her research interests include human-computer interaction (specifically user-centered design), and the appropriateness of various evaluation techniques for ubicomp. She is a member of the ACM and ACM SigCHI. Contact Consolvo at Intel Research Seattle, 1100 NE 45th St., 6th Floor, Seattle, WA 98105; sunny@intel-research.net.

Miriam Walker is a senior research analyst at Peace Software. Her research interests include assistive technology, research methodologies, and compatibility of software engineering and usability practices. She received an MS in computer science from the University of California, Berkeley. She is a member of ACM SigCHI, ACM SigCHINZ, and the UPA. Contact Walker at PO Box 37-580, Parnell, Auckland, New Zealand; miriam.walker@peace.com.

REFERENCES

1. R. Want et al., "The Personal Server: Changing the Way We Think about Ubiquitous Computing," *Proc. 4th Int'l Conf. Ubiquitous Computing*, IEEE CS Press, 2002, pp. 194–209.
2. S. E. Hormuth, "The Sampling of Experiences in Situ," *J. Personality*, vol. 54, no. 1, Mar. 1986, pp. 262–293.
3. M. Csikszentmihalyi and R. Larson, "Validity and Reliability of the Experience-Sampling Method," *J. Nervous and Mental Disease*, no. 175, 1987, pp. 526–536.
4. S.E. Hudson et al., "Predicting Human Interruptibility with Sensors: A Wizard of Oz Feasibility Study," *Proc. Conf. Human Factors and Computing Systems*, ACM Press, vol. 5, no. 1, Apr. 2003, pp. 257–264.
5. L.F. Barrett and D.J. Barrett, "An Introduction to Computerized Experience Sampling in Psychology," *Social Science Computer Rev.*, vol. 19, no. 2, Summer 2001, pp. 175–185.
6. L. Palen and M. Salzman, "Voice-Mail Diary Studies for Naturalistic Data Capture under Mobile Conditions," *Proc. Conf. Computer Supported Cooperative Work*, 2002.

For more information on this or any other computing topic, please visit our Digital Library at http://computer.org/publications/dlib.