

# Scavenger Hunt: An Empirical Method for Mobile Collaborative Problem-Solving

*The scavenger hunt prototyping model avoids some challenges of complex field studies, while supporting significant mobility and collaboration testing in a more controlled environment.*

Pervasive computing enables field researchers to accomplish efficient fieldwork in teams that they once performed in isolation, over several trips, or not at all. Field researchers can deduce new information from findings they make while in the field, and apply it immediately to the situation at hand. This is especially important in fields where the time or resources to conduct several studies isn't available. This domain can be termed *mobile collaborative problem-solving*.

Because collaboration systems that support mobile users can be costly and difficult to create, designers need tools to ensure later iterations of their systems will function as expected. Mobile collaborative problem-solving is a young domain of study, and not much is known about how people can collaborate effectively in the field using computers. Moreover, applica-

tions that are found to be effective in a laboratory study might not actually be effective in the setting where they will be used (for example, laptops for use by patrolling police officers). Evaluating mobile collaborative systems requires methods for studying team use of these systems in realistic yet controlled settings. Before deploying a mobile collaborative problem-solving system, early evaluation methods can help identify problem areas in the user experience. By troubleshooting design problems while the product is still in development, designers can save time and money.

We suggest a *scavenger hunt* model that provides a viable prototyping method. A scavenger hunt is a typical mobile activity that both adults and children can perform. Participants are divided into teams and given a list of items, often unrelated and obscure. The first team to collect all the listed items within a given time limit wins the game. We use the essential elements of this game format—a timed task, teamwork, and mobility—to create a prototyping method for mobile collaborative problem-solving systems. These elements of the scavenger hunt mimic several field challenges in the lab. For example, a search for a missing person, an educational field trip to a nature preserve, or repeated trips to the same area for field research all share these characteristics. In these situations, the data capture and the data analysis feed into each other immediately. This tight coupling of data capture and analysis is useful in other situations as well.

The scavenger hunt empirical tool lets programmers and system designers study the effectiveness of their mobile collaborative problem-solving environments in a setting that offers laboratory-like controls while mimicking the real-world problems facing mobile users (see the “Finding a Medium between Laboratory and Field Studies” sidebar for some other work in this area). It presents users a well-defined problem that only the group can solve and simultaneously requires them to navigate a public area. By observing participants during a scavenger hunt trial, we learn more about the field problems they'll encounter regarding software, group dynamics, infrastructure, and mobility.

Michael Massimi  
University of Toronto

Craig H. Ganoe  
and John M. Carroll  
Pennsylvania State University

## Finding a Medium between Laboratory and Field Studies

Researchers have explored both the use and usefulness of laboratory versus field studies in mobile human-computer interaction. Jesper Kjeldskov and Connor Graham examined 102 research papers in mobile HCI published between 2000 and 2002 and found that only 41 percent involved some kind of user evaluation (71 percent laboratory, 19 percent field, and 10 percent through survey).<sup>1</sup> At least in some cases, this tendency toward laboratory studies appears to be justified.

Kjeldskov and his colleagues compared laboratory and field evaluations of their MobileWard prototype to support hospital morning procedures for nurses.<sup>2</sup> Of 37 usability problems identified in the study, 36 occurred in the laboratory setting while only 23 occurred in the field. Researchers spent 34 staff hours on the laboratory evaluation of six users and 65 staff hours on the field evaluation of another six users. Although this might seem to be a strong argument for conducting most mobile-application evaluations in the lab, the application being evaluated was for a single user and geared toward data collection and retrieval.

Melanie Kellar and her colleagues designed and conducted a field study in which pairs of participants collaborated with their software in the scavenger-hunt-like City Chase ([www.thecitychase.com](http://www.thecitychase.com)).<sup>3</sup> They argue that their field study let them observe six external factors that are difficult to control and that impact both research into and adoption of mobile technologies:

- *software*—failures of the software being tested and wireless connectivity failures;
- *materials*—lack of a home base for items such as paperwork and equipment;
- *social considerations*—influence by the public (interactions and self-consciousness);
- *weather/environment*—rain, wind, and sun (one study noted, “tree sap dripped onto equipment”);

- *audio and video*—background noise and shaky and poor video angles; and
- *mobility*—observation and note-taking difficulties in crowded areas, while crossing streets, and while moving in general.

Although we agree that these factors come into play in field-study settings, whether most of them would have much influence on software user interface design (rather than the ability to collect field data) is less clear.

Khai Truong and his colleagues make a strong case for prototyping mobile computing systems with users in mind.<sup>4</sup> Too often, they claim, programming tools and systems are “device-centric, rather than user-centric.” We believe that the scavenger hunt can be used in two ways: as a tool to assure that users will find mobile collaboration systems useful and as a lens for studying individual and group planning. We attempt to look beyond the device and offer a method for examining how people plan and act in the field.

### REFERENCES

1. J. Kjeldskov and C. Graham, “A Review of Mobile HCI Research Methods,” Proc. Int’l Conf. Human Computer Interaction with Mobile Devices and Services (Mobile HCI), ACM Press, 2003, pp. 317–335.
2. J. Kjeldskov et al., “Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field,” Proc. Int’l Conf. Human Computer Interaction with Mobile Devices and Services (Mobile HCI), ACM Press, 2004, pp. 61–73.
3. M. Kellar et al., “It’s a Jungle Out There: Practical Considerations for Evaluation in the City,” Proc. Conf. Human Factors in Computing Systems (CHI 05), ACM Press, 2005, pp. 1533–1536.
4. K.N. Truong et al., “How Do Users Think about Ubiquitous Computing?” Proc. Conf. Human Factors in Computing Systems (CHI 04), ACM Press, 2004, pp. 1317–1320.

### Interactive prototyping in mobile collaborative environments

Linchuan Liu and Peter Khooshabeh describe the advantages and disadvantages of paper versus interactive-prototyping techniques.<sup>1</sup> Specifically, they suggest that *fidelity* (look and feel) and *automation* (amount of human intervention) are important dimensions for gauging a prototyping technique’s success. Our

scavenger hunt prototyping and empirical analysis technique offers excellent fidelity but only moderate automation.

Liu and Khooshabeh also argue that interactive prototypes are necessary parts of the design process and that interactive-prototyping methodologies are important for the product’s success. Furthermore, they argue that “[a]lthough prototyping has been used with great success in obtaining usability data during the design

of traditional UIs, its use in ubicomp has not been thoroughly investigated.”<sup>1</sup> As an empirical tool, the scavenger hunt methodology advances progress toward a realistic prediction model about the success of a mobile collaborative problem-solving environment.

Previous work has explored the individual areas of mobile collaboration and collaborative problem-solving. Their fusion, however, requires designers to

rethink their strategies when creating software for field users. Because problem solving is the overarching goal, designers must equip systems with rapid data-acquisition techniques, easy data retrieval, and unobtrusive analysis tools. Users should be able to spend their time piecing together information rather than navigating an interface. They should be able to easily share the data or hide it at their discretion. At the same time, the design must consider user mobility; users will be traveling through an environment that also requires monitoring and an occasional response.

### **A blog-based mobile collaborative problem-solving system**

Evaluating the scavenger hunt methodology required developing software that was representative of current trends in information sharing. The ultimate goal was to use a scavenger hunt to unearth the software's design flaws. As experimenters, we weren't evaluating the users or even necessarily the product, but rather how well the scavenger hunt format highlighted the problems our users faced. On the other hand, designers using this tool will be more interested in the flaws the scavenger hunt uncovers than the method itself.

A rapid examination of current information-sharing tools led us to create a blog-style collaboration and problem-solving system. Many organizations use blogs to distribute information to interested parties. For field agents working on a project over an extended period, blogs offer a chronological way to structure data and findings. They're also relatively easy to program and maintain. This approach also expedited data collection because blogging software already contains much of the necessary user information (posting time, name of the user submitting the post, and so on).

Our blog was simple because we didn't want users to become too involved in understanding its features during the user studies. Users added posts using a small onscreen (soft) keyboard on the handheld computer. New posts were appended to the end of the blog. However, users could rearrange a post's position relative to other posts by pressing "up"

**Because problem solving is the overarching goal, designers must equip systems with rapid data-acquisition techniques, easy data retrieval, and unobtrusive analysis tools.**

or "down" buttons. They could also edit their posts in place (that is, they didn't need to submit an edit request). After editing, users could save their changes by pressing a "save" button next to the post. Finally, users had to request updates to the blog manually—pressing a "refresh" button updated the blog to the most current version.

### **Empirical environment**

An accurate methodology for studying the usability of collaborative tools in pervasive environments will give participants a realistic representation of a collaborative task. To be as realistic as possible, the methodology must support several requirements:

- The task is well defined. It doesn't inundate users with information, nor is its premise so scant that users can't adequately create a plan of action. The problems in our scenarios have a specific goal.
- The environment allows participant monitoring and rapid data collection.
- The task is important enough to participants that they won't get discouraged and disengage from the task.
- The task requires both individual and

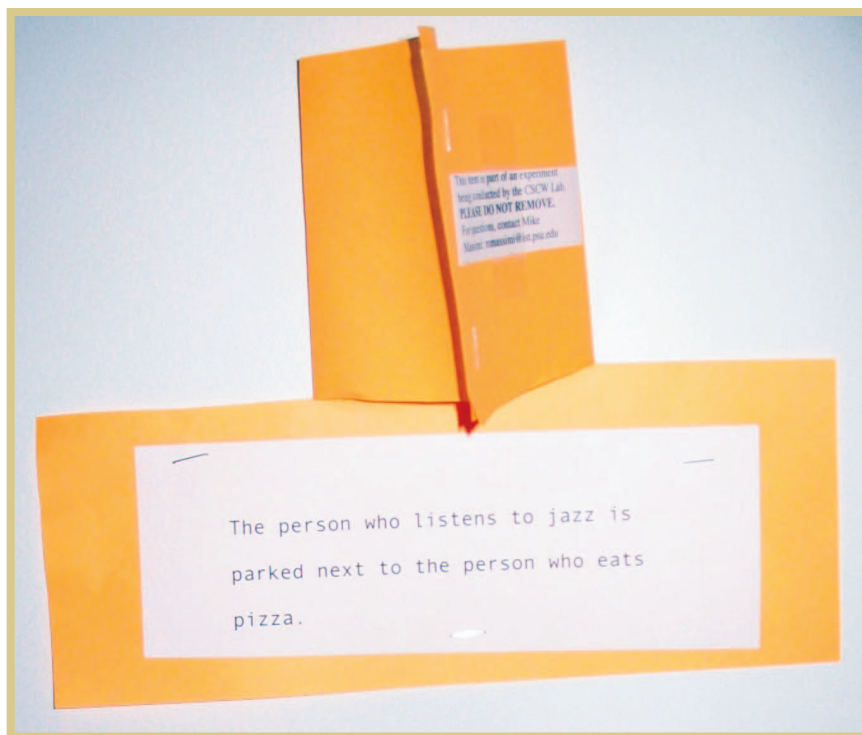
collaborative work within the problem-solving system.

- Participants must have a limited amount of time in which to complete the task. Given unlimited time, participants might not exploit the tools to their fullest potential.
- The pieces of information that users manipulate must be independently

meaningful, but collectively powerful enough to accomplish the goal. Independent meaning is necessary so that the users can understand the relationships between pieces of information.

- The distribution of the pieces of information has no predefined logic—that is, no one person knows everything. Users should acquire the pieces of information semirandomly before sharing them.
- No piece of information deductively implies another piece. So, users can't complete the goal simply by discovering a single piece of knowledge. Instead, they must compile and analyze a constellation of facts.

In our scavenger hunt method, participants search for clues that form the basis of a logic puzzle. The clues are located throughout an academic building. Each participant carries a handheld computer with wireless Internet access and a walkie-talkie as well as a slip of paper with a transliterated version of Einstein's famous riddle (we altered the riddle to prevent participants who might have previously encountered it from immediately recognizing it). The riddle's premise is as follows:



**Figure 1.** A sample clue. Clues are mounted to walls and tables throughout the building.

*In a parking lot, five cars of different models are parked next to each other. The owner of each car has a different profession. The five car owners each play a different sport, listen to a different type of music, and eat a different food. The question: Who eats lo mein?*

Participants have one hour to walk through the building looking for clues. On finding a clue, participants use their handheld computer to record the clue and broadcast it to their teammates. A Web browser on the handheld computer displays the blog containing all the clues. Once participants feel they have enough clues to solve the riddle, they return to the starting point and verbally report their answer to the experimenter.

### User study parameters

All study participants were students (ages 15 through 17) involved in a summer program for gifted youth. They participated in the study as an extracurricular activity. Although participants stated that they didn't have extensive experience with mobile devices, they rapidly adjusted to the input methods of the handheld computers we gave them. We conducted four

trials with different groups of three students each. The first trial was a pilot study, and the other three were full studies.

We placed 24 clues on brightly colored paper (see figure 1) and distributed them strategically throughout an academic building with 7,700 square feet of public space on three floors. Participants could easily access all clues on foot. In each trial, participants needed 18 of the clues to solve the puzzle. The remaining six clues served as distractors, which ensured that the participants were actually analyzing the information they collected, as opposed to simply copying it and deferring analysis until later. In the pilot trial, two of the distractors were irrelevant—they provided information that participants didn't need to answer the riddle. Another two were premises—the actual riddle described earlier. The remaining two distractors were duplicates—they were simply copies of clues that were elsewhere in the building. In subsequent full trials, we replaced the irrelevant clues and premises with four more duplicates to make the task easier to complete.

We gave each participant a Hewlett

Packard iPAQ h5450 handheld computer running Microsoft Windows Mobile. An integrated IEEE 802.11b adapter connected the handheld computers to the building-wide wireless network via a virtual-private-network client. We wrote the blog software in PHP and used a MySQL backend database to collect data and represent clues. The Internet browser was Microsoft Internet Explorer. The blog software lets users add, delete, edit, and move posts up or down on the Web page, and tracks these changes.

Camera operators followed one or two participants in each group, recording their walks through the building. We asked all participants to think aloud as they worked with their handheld computers so that we could better gauge their reactions to the task. At the end of the task, we videotaped interviews with the participants. Participants then completed questionnaires containing 24 items with a seven-point Likert-type response measure and three open-ended questions.

### Trial breakdown

The pilot group consisted of three male participants. As mentioned earlier, this group encountered irrelevant clues as distractors, which we later replaced with duplicates. Furthermore, we didn't give this group the premise at the outset; instead, we disguised it as an additional clue in the field, as we discuss in more detail later. The group commented on the lack of automatic refreshing, the interface's slow scrolling speed, and general frustration with the software. They divided the building's floors among themselves, assigning one person to each floor. The participant who discovered and entered a clue was accountable for understanding its contents and rapidly responding to teammates' questions about the clue's content using walkie-talkies. The group also delegated particular tasks to individual team members.

As a direct result of our observation of participants during the pilot study, we modified the empirical method for the subsequent three full study groups. At the beginning of the pilot study experiment, we simply asked the participants “Who eats lo mein?” and let them determine the rest of the riddle on their own. Their responses in the questionnaire showed that they found the task very difficult. This indicated that to make immediate productive use of a handheld computer in a mobile and collaborative environment, users must have a cognitive scaffolding of sorts—that is, a clear goal and well-understood instructions. Similarly, well-structured tasks can make better use of handheld computers than ill-defined ones.

The first group in the full study consisted of only two people, as one participant failed to arrive. Both participants were female. One participant seemed especially dispassionate about the task but also exhibited a knowledge of which clues she had encountered earlier in the task. She simply skipped over these clues and rapidly discarded distractors. However, she was not thorough in exploring the building and therefore her team failed to find several clues. They didn’t split up the building in terms of work, but instead wandered individually with little communication as to who would be responsible for which clue. They moved clues that had relevance to one another closer together on the Web blog as well.

The second full group consisted of three male participants. They immediately divided the building among themselves and began collecting clues on their assigned floors. They created an aggregate post that served as a “good copy” of their knowledge thus far. They attempted to draw a picture in one of the blog posts using textual symbols as a substitute for lines (that is, ASCII art), but seemed to disregard it after a time as useless. In this group, one of the partic-

ipants seemed to assume a leadership role. This participant began to assign tasks to the other two team members, while he maintained the good-copy post and thought aloud the most.

Like the second group, the third and final full group consisted of three male

participants. This group also split the floors among themselves. They immediately recognized the type of problem when they were given the premise. After collecting most of the clues, this group accidentally deleted the post that contained their aggregated knowledge. To remedy this, they sent one person out to gather all the information again while the other two tried their best to complete the puzzle.

From these observations, we noticed particular flaws in our blogging software that we would have otherwise likely disregarded. The accidental deletion of clues is an excellent example. As observers, we saw this as a red flag—we needed to design a way for the users to retrieve lost posts.

### Results and user responses

Of the three full studies and one pilot study, only the second group successfully answered the overall riddle. Table 1 shows responses to the postexercise questionnaire.

Participants’ responses to the questionnaire make clear the flaws in our blogging software. Items with responses near the extremes (+3 or -3) indicate areas that are especially important to address. Although we’ve only used this tool with a single product (a custom

blog), our run indicates that the tool was successful in uncovering design flaws that might otherwise have gone unnoticed until deployment. We look forward to trying the scavenger hunt prototyping tool with other products to determine its flexibility as a method.

## To make immediate use of a handheld computer in a mobile, collaborative environment, users must have a cognitive scaffolding of sorts—a clear goal and well-understood instructions.

The data gathered from the scavenger hunt on how to improve our blog came primarily in three forms:

- The *questionnaires* let us maintain consistency across trials and provided background information on our participants in a formal manner.
- The experimenters’ *observations* let us monitor users’ progress at any given point in the run. Observations also helped us to understand subtleties in the design that caused emotional responses (such as anger or frustration). By observing users in a constrained, realistic environment, we can capture items that aren’t readily apparent in questionnaires or interviews.
- Semistructured *interviews*, however, let us direct our line of inquiry at the end of each session on the basis of what had occurred in that session. This let us tease out responses from the participants that, again, weren’t readily observable in a questionnaire. Participants also suggested valuable new features or changes to our blogging software that we hadn’t considered. Because we conducted group interviews, participants could build on each others’ suggestions for improvements, leading to more specific or alternative designs.

Table 1. Postexercise questionnaire responses. Ratings range from -3 (strongly disagree) to +3 (strongly agree). Average response calculated from eight participants in three full studies (pilot study data not included).

| Number | Question text   | Average response |
|--------|---|------------------|
| 1      | Prior to today, I had extensive experience with handheld computers.   | -1.625           |
| 2      | Prior to today, I had extensive experience with mobile telephones (cell phones).                                    | -1.625           |
| 3      | I found it relatively easy to enter clue information into my handheld computer.                                     | 0.625            |
| 4      | I felt I was able to control the sharing and receiving of clues efficiently.  | -0.125           |
| 5      | At all times, I was generally aware of the location of each of my teammates.  | 1.125            |
| 6      | Most of the time I was unsure what each of my teammates was doing.  | 0                |
| 7      | I became frustrated because I worked on a task and later found out that someone else had already completed it.      | -0.5             |
| 8      | I found it easy to rearrange the order of the clues as displayed on my handheld.                                    | -2               |
| 9      | I became frustrated trying to share clues with my teammates.  | 0.625            |
| 10     | I thought this task was difficult to complete.  | 1                |
| 11     | I thought the handheld computer helped me organize information effectively.   | -1.75            |
| 12     | I would have liked a space to work on the puzzle that was not visible by my teammates.                              | -0.875           |
| 13     | The handheld computer distracted me from other tasks.   | -1.5             |
| 14     | I would have liked to draw pictures during the task.  | 2.75             |
| 15     | I would have liked to make a table during the task.   | 3                |
| 16     | I would probably only use a handheld computer if I were working alone.  | 0                |
| 17     | I would not want to use a handheld computer for group work at school.   | -1               |
| 18     | I feel all team members contributed equally to the completion of the task.  | 2.75             |
| 19     | If I were to complete this task again, I would want to use a handheld computer with the same software I used today. | -2.25            |
| 20     | If I were to complete this task again, I would want to use a handheld computer more than paper and pencil.          | -1.125           |
| 21     | If I were to complete this task again, I would want to use a mobile phone more than a handheld computer.            | -0.5             |

The scavenger hunt method's value lies in its ability to compromise between several extremes. It gives experimenters a good amount of control, but not so much that users are guided to an outcome. Users instead are free to push the software to the boundaries of its use. Furthermore, the scavenger hunt method doesn't require a trip to the target locale where the software is to be deployed. However, it still offers a fair amount of ecological validity by mimicking the situations faced by workers in the field.

## Discussion

Helen Cole and Danaë Stanton offer insight from three case studies that leverage mobile devices for collaboration: KidStory, Hunting of the Snark, and Ambient Wood.<sup>2</sup> Because, as they note, people "use their own mobility and the mobility of artifacts to coordinate their

collaboration with one another," it is important to investigate how mobility impacts collaborative problem-solving. Their work does not focus on problem solving specifically, but instead examines collaboration in adjacent areas: learning environments (Ambient Wood), children's entertainment (Hunting of the Snark), and storytelling (KidStory). Their use of location-based information in Ambient Wood and Hunting of the Snark parallels our own use of clues in the scavenger hunt.

As in Hunting of the Snark, we modeled our prototyping tool as a game. This lets us leverage people's preexisting knowledge about games. They know that the task is timed and that they need to overcome an obstacle to achieve a goal. Because games are familiar and fun, users feel more comfortable than when in a stilted, strictly controlled lab-

oratory environment. The sense of challenge also encourages participants to remain engaged and active. They are motivated to win throughout the trial.

Because the scavenger hunt acts as an early-stage evaluation tool for iterating through designs of mobile collaborative problem-solving software, developers would likely have to run multiple scavenger hunts each time they change the software. Changes suggested by scavenger hunts will influence the product's design in future iterations. Eventually, scavenger hunts will fail to reveal new flaws in the software. Then, developers will need to perform actual field tests or deploy the software. In either case, we hope they can benefit from the software tests and debugging in a realistic environment.

Although the scavenger hunt method effectively mimics the challenges of working in the field, it is by no means perfect.

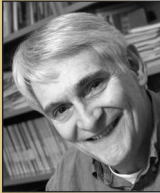
## the AUTHORS



**Michael Massimi** is a graduate student with the Dynamic Graphics Project in the Department of Computer Science at the University of Toronto. His research interests include assistive technologies for the cognitively impaired, computer-supported collaborative work, ubiquitous computing, and human-computer interaction. He has a BS in computer science from the College of New Jersey. He's a student member of IEEE and the ACM. Contact him at [mikem@dgp.toronto.edu](mailto:mikem@dgp.toronto.edu).



**Craig H. Ganoe** is a senior research associate with the College of Information Sciences and Technology at Pennsylvania State University. His research interests include human-computer interaction, in particular applied to mobile computer-supported cooperative work, community computing, and collaborative learning. He has an MS in computer science from Virginia Tech. He's the information director for the *ACM Transactions on Computer-Human Interaction* and a member of the ACM. Contact him at [cganoe@ist.psu.edu](mailto:cganoe@ist.psu.edu).



**John M. Carroll** is the Edward M. Frymoyer Chair Professor of Information Sciences and Technology at Pennsylvania State University. His research interests include methods and theory in human-computer interaction, particularly as applied to networking tools for collaborative learning and problem solving, and the design of interactive information systems. He serves on several editorial boards for journals, handbooks, and series and is editor in chief of the *ACM Transactions on Computer-Human Interactions*. He is a fellow of the ACM, the IEEE, and the Human Factors and Ergonomics Society. Contact him at [jcarroll@ist.psu.edu](mailto:jcarroll@ist.psu.edu).

Its use can mean settling for a watered-down version of the actual task to be completed. Usability flaws are always possible if the circumstances aren't identical. What's more, in the laboratory you can't compensate for environmental problems resulting from working in the field. Users might be confused by the task or react negatively to the time limit or competitiveness that can arise in the game-like circumstances. The scavenger hunt, like any laboratory-based simulation of real-life conditions, doesn't perfectly mimic the realities workers in the field face. However, its low overhead and realistic setting makes it an attractive substitute for field conditions.

The scavenger hunt is a lightweight, cost-effective way to prototype a project while maintaining similar conditions to those faced in the field. Further work in this area will strive to improve the collaborative problem-solving software for mobile users and to refine the scavenger hunt model for evaluating design. Our method improved rapidly from initial experimental design, to the pilot study,

to our user studies. We believe that more user studies will help us refine the empirical method. We're interested in hearing from practitioners and designers who use scavenger hunts to test their products and hope we've provided a foundation that developers can build on to create better prototyping tools. ■

## ACKNOWLEDGMENTS

We thank Matthew Dalius, Michael Race, and Brent Schooley for technical assistance; Janet Montgomery for testing the riddle and feedback; Matthew Peters and Cecelia Merkel for recruitment assistance; and the members of Penn State's Computer-Supported Collaboration and Learning lab for their support and accommodation.

## REFERENCES

1. L. Liu and P. Khooshabeh, "Paper or Interactive? A Study of Prototyping Techniques for Ubiquitous Computing Environments," *Proc. Conf. Human Factors in Computing Systems (CHI)*, ACM Press, 2003, pp. 1030-1031.
2. H. Cole and D. Stanton, "Designing Mobile Technologies to Support Co-Present Collaboration," *Personal and Ubiquitous Computing*, vol. 7, no. 6, 2003, pp. 365-371.

# Tried any new gadgets lately?

Any products your peers should know about? Write a review for *IEEE Pervasive Computing*, and tell us why you were impressed. Our New Products department features reviews of the latest components, devices, tools, and other ubiquitous computing gadgets on the market.

Send your reviews and recommendations to [pvcproducts@computer.org](mailto:pvcproducts@computer.org) today!



[www.computer.org/pervasive](http://www.computer.org/pervasive)