Reprint of paper delivered at
UPA 2005, sponsored by the
Usability Professionals' Association

Technology and Techniques for Conducting Instant-Messaging Studies

Deborah Hinderer Sova
Tec-Ed, Inc.
deborah@teced.com

Laurie E. Kantner
Tec-Ed, Inc.
laurie@teced.com

ABSTRACT

Designing usability lab tests of instant-messaging services, whether conducted on a hand-held device or a computer, presents unique challenges for the testing team. This presentation describes three instant-messaging studies and the technology and techniques used to instill realism and maintain rigor.

INTRODUCTION

Instant messaging, or IM, is a popular consumer and business tool that offers "real-time synchronous collaboration. Real-time collaboration tools allow geographically dispersed people to "meet online" [1]. Now, thanks to easy (often free) access to IM software and the proliferation of mobile phones, as well as recent strides in the capabilities of both, a whole new IM "patois" has emerged, making IM hugely popular with younger adults leading mobile lifestyles.

"IMing" also has become a desirable way for computer users—both personal and business—to get quick responses [2] to messages, knowing when their colleagues or friends are online and available to converse. In July 2000, there were an estimated over 100 million IM users, over 12 million of them business IM users [2]; we can only assume those numbers have grown in the last five years.

The authors have designed and conducted lab studies of hand-held and computer instant-messaging services over the past 10 years, beginning with two-way pagers used mainly by early adopters. Both authors are senior-level usability researchers who have conducted scores of studies—including over a dozen mobile and/or instant-messaging studies—with hundreds of users.

This paper describes three of the authors' recent lab studies of instant-messaging and video instant-messaging services and the technology and techniques used to instill realism and maintain rigor. Note that we will not discuss the merits of specific brands of equipment or IM software. Rather, our paper focuses on ways to plan and set up a lab study for conducting IM tests using realistic scenarios.

CHALLENGES OF TESTING REAL-TIME COMMUNICATION REALISTICALLY

Designing usability lab tests of instant-messaging services, whether conducted on a hand-held device, a computer, or both, presents unique challenges for the testing team. At a high-level, testing real-time IM communication is unlike testing "traditional" websites or software, because IM studies require that test users and test facilitators be interactive with each other, not just the computer. In IM studies, we need to evaluate not only users interacting with a computer, but also the reactions that users have with their IM partners [1]. Test facilitators have to be able and willing to take on the roles of facilitator, observer, *and* interactive participant. We also have to devise scenarios that are realistic enough to motivate people to engage in a conversation and be comfortable doing so.

At a practical level, testing real-time IM communication—from two-way pagers 10 years ago to mobile phones and PDAs today—means planning for special technology and lab setups that often require:

- Integrating a hand-held IM device with computer software
- Having to video-record both the computer screen and the device during the session
- Finding ways to observe user interaction with the small device interface without infringing on the user's personal space
- Dealing with the mobility of a hand-held device and having to videorecord a sometimes "moving target"
- Depending on sometimes unreliable wireless networks in out-of-service range locations ("Can you hear me now?" [3])

TWO-WAY PAGER AND PC—LEGACY CASE STUDY

The authors conducted a usability test of an early two-way paging system, including the ability to send messages using the device, using a phone, and using a PC, and the ability to reply to messages. We provide this case history as an interesting glimpse of times past—over a decade ago—from which the lessons learned helped us design our more recent messaging studies.

Special Challenges

In addition to the challenges listed above under "Challenges of Testing Real-Time Communication Realistically," this early study posed a recruiting challenge that we don't find as difficult today. While one-way paging (using "beepers") was quite common at the time, two-way paging was much less so. It was difficult to recruit adequate numbers of two-way paging customers—there were only two main manufacturers of these devices. So, we had to recruit beeper users who responded positively to the concept of "two-way" paging—responding to an incoming message without having to find a phone.

Test Team

In our pager studies, we used a team of two—a test facilitator/note-taking observer and a second note-taking observer who also acted as real-time messaging partner for the test participant. The test facilitator sat in the same room as the participant, and the messaging partner sat in an observation room where she could follow the session.

Technology and Equipment Setup

The participant sat in a usability lab; a wall-mounted video camera, controlled by the note-taker in the observation room, captured the participant's actions with the pager and then on a PC. The participant also worked with a telephone during the study, and the movable camera captured that interaction as well. Three pagers were used: the participant's pager, a pager for the participant's messaging partner, and a spare pager as backup. The observer sent messages and responses to the participant via a two-way pager or via the paging service's telephone operator.

Scenarios

The two-way pager study took place in the days before everyone carried a mobile phone. The typical use assumed in the scenarios created for the study was similar to the use today of mobile phones, IM messaging, text messaging, and email. Tasks included the participant receiving work-related and personal messages to which they could either choose a "canned" reply from the unit or enter their reply by selecting letters from a screen. Scenarios included a notification of meeting location change, a request for a dinner appointment, and scheduling a meeting.

Conclusions

The study confirmed that two user researchers were sufficient to conduct a usability study of two-way communication that involved multiple sending devices. We learned that marking an area on the table or desk for use of a hand-held device facilitated recording user actions without impairing natural use of the device. Our study participants confirmed for us in a post-test questionnaire that the tasks we asked them to perform, as well as the scenarios we suggested, were realistic. We also *observed* that they did not hesitate to engage in the scenarios we devised.

We applied the lessons we learned about testing devices from our pager studies to our current studies. Fortunately, we now have affordable, digital recording equipment and higher-quality monitors, so we can capture and observe participant interaction with a small-device interface more easily and clearly than we did 10 years ago. However, updated equipment alone does not make a real-time communication study successful, as we discuss under "Two Mobile-Phone and PC Case Studies," below.

TWO MOBILE-PHONE AND PC CASE STUDIES

The authors each conducted a study that tested the use of instant messaging on a mobile phone. The first study also tested the usability of a web application for installing the application on the phone. The second study focused strictly on the phone IM application. The participants targeted for both studies were mobile-phone and IM users, 18 to 26 years old, who had higher-end phones with text-messaging capability.

Special Challenges

In addition to the challenges listed above under "Challenges of Testing Real-Time Communication Realistically," these IM/mobile-phone studies posed new issues. Not only did we need to observe users interacting with a small screen, but we also needed to see and record what buttons and numbers they pressed—many more than were available on a two-way pager—to converse with an IM partner. In the first study, we also had to change our session approach in response to unexpected participant reactions to conversing with an unknown partner, as described under "Test Team," below.

In addition, we had to create realistic "buddy lists," which held many more entries than the address books of our two-way pagers. We used the IM software to create unique user IDs for each test participant, the facilitator, and the designated IM partner, as well as other plausible "IM-buddy-like" names, such as "Palomine," "Chum789," "Daisy Lady," and "Debzupa."

Another challenge we faced was that the test facilitator had to use more than one ID at the same time to be able to initiate and participate in multiple conversations, which wasn't supported by the IM software. We worked around this issue by using the Web version of the IM application with one ID and the Windows version with a second ID.

Test Team

In the first of the two IM/mobile-phone studies, the original plan was to have a two-person team: one person who facilitated the sessions, and a second person who served as messaging partner. The messaging partner's job was to initiate IM conversations and then keep them active. (To instill realism, the messaging partner was not in the test room, and received instruction from the test facilitator via IM communication to initiate the next planned conversation.)

After two sessions, the test facilitator changed the approach and took on the role of messaging partner. The decision to change the protocol was based on observing participants' discomfort at conversing with someone they both did not know and could not see (their discomfort was conveyed through body language and the tendency to immediately end conversations). When the test facilitator began serving as messaging partner, the participants visibly relaxed and continued conversing via IM.

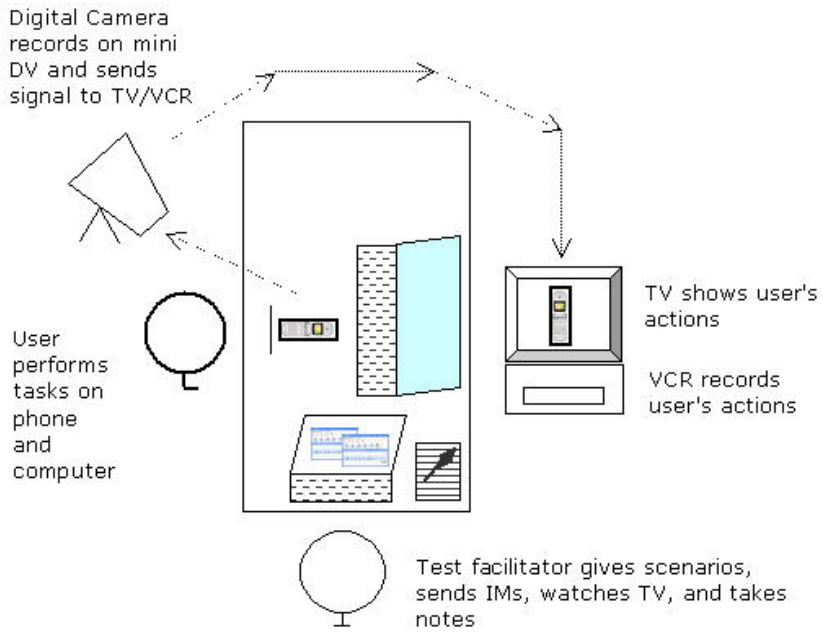
Because this second approach worked so well in the first study, both for participant comfort level and for reducing team coordination and communication during sessions, we planned for the test facilitator to act as the sole messaging partner as we designed the second study.

Technology and Equipment Setup

For the first study, the participant sat at a table, with a mobile phone placed in front of him or her and a PC connected to the Internet within reach. Based on our experience from the pager studies, we taped an area on the desk in front of the participant and requested that s/he keep the mobile phone within the taped boundaries so our camera would be able to capture all interaction.

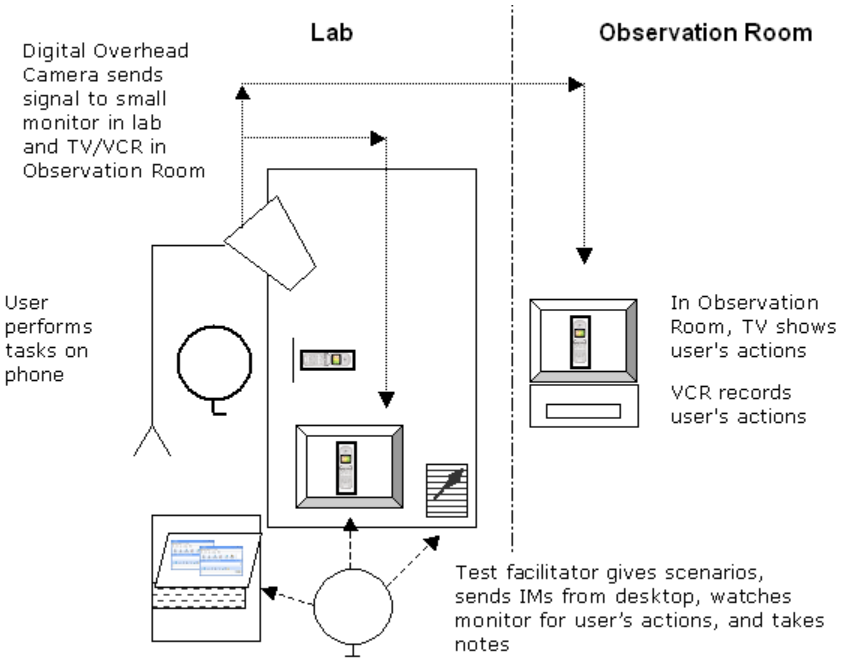
To solve the issue of more easily observing participant interaction on a small device screen, and the device's many more buttons and keys than were available on the two-way pagers—without encroaching on the user's personal space—we planned to have the facilitator observe the interaction through a TV monitor, in much larger-than-actual size. To accomplish this, we aimed a professional-quality digital video camera over the participant's shoulder at the mobile-phone screen. The signal from the digital camera went to both the digital recording tape and through a VCR to a TV monitor, placed where the test facilitator could view it and see the participant's actions (as well as make sure the recording was capturing the phone screen). In addition, the test facilitator used a laptop connected to the Internet to send messages. (See Figure 1, below.)

Figure 1: Setup of First Mobile Phone Instant-Messaging Study



For the second study, the participant did not use a PC (we tested only the mobile-phone interface), and the professional-quality digital video camera our client supplied was mounted on an overhead camera boom, aimed directly at the mobile-phone screen from above. The digital camera sent a signal to a small monitor from which the facilitator could readily view the participant's interaction with the mobile phone (held within a taped-off area of the table), as well as to a TV and VCR recorder in the observation room. The facilitator sent messages to the participant from a PC on an adjacent small table. (See Figure 2, below.)

Figure 2: Setup of Second Mobile Phone Instant-Messaging Study



Scenarios

For both studies, the sessions consisted of typical IM tasks that would explore participant's use of the phone-based IM user interface, especially multiple conversations, perceptions of IM service status when the phone was turned off, and anticipated uses of the feature. The scenarios introduced conversation topics that were immediate but not too personal (to respect the actual interpersonal situation between participant and facilitator), and that enabled a telescopic style of communication common to today's IM users—an etiquette characteristic that hasn't changed since the days of two-way pagers.

As noted above, the messaging partner/test facilitator had two identities during the session and sent first under one identity using the Web application, and then under the second using the Windows version of the application. Conversations ranged from topical (asking if the participant saw a recent local-rivals college sports match) to logistical (asking the participant to meet at a bookstore at a particular time). Two conversations supposedly with different partners addressed the study goal of learning how easily the participant could navigate between conversations.

Conclusions

Pilot testing of the logistics of real-time conversation studies is important to ensure that the equipment supports the scenarios. For example, the phone service being used by the mobile phone had good reception at the user researcher's desk during the study planning, but poor reception in the usability lab. In addition, being able to maintain two separate IM logins on one computer is not feasible with most of today's IM programs.

We also observed that the study participants readily embraced the tasks and scenarios, once they felt comfortable with whom they were IMing.

VIDEO/PC INSTANT-MESSAGING CASE STUDY

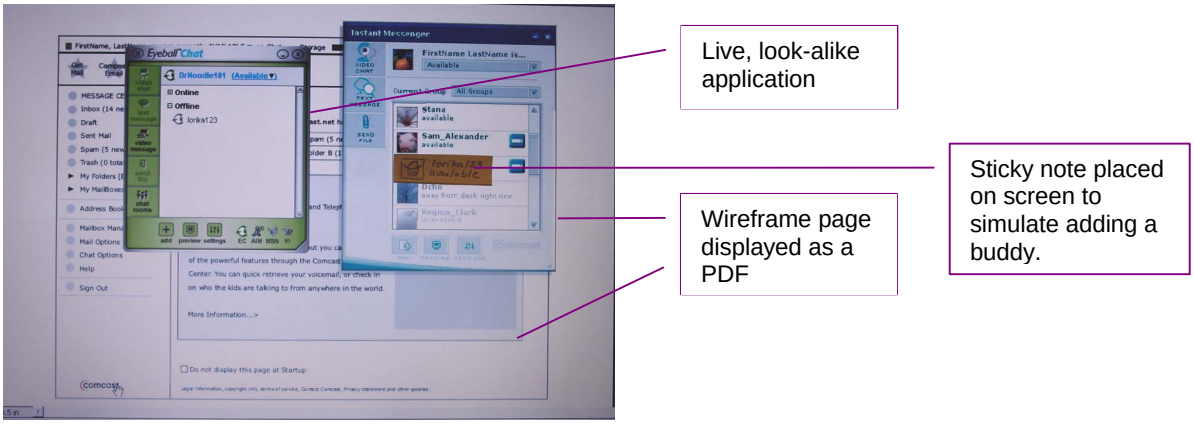
This case history describes usability testing of video instant-messaging software on a computer. We tested participants who were email users with high-speed Internet connections (needed for optimal video performance). The majority of the participants also had IM experience, and some had previously used a video application such as, video mail, video IM, or video conferencing. We tested installation and first use of the video IM system.

Special Challenges

This study was more different from the two-way pager studies than the mobile-phone IM studies were, and therefore, had some additional challenges, including:

- Devising scenarios where sending and receiving video instant messages—a new technology for most people—would be a plausible means of communicating with someone. We successfully created a "best friend moving to another town for a new job" whom the participant would miss seeing.
- Dealing with new technology issues such as audio echo looping and video lag, which, unfortunately were inherent in the new technology. We could mitigate echo looping slightly, by keeping the computer's speakers pointed away form the webcam's microphone (standard good audio practice) but could not help the video lag.
- Overcoming firewall issues and other internal network privacy issues to allow internal communication (from lab to observation room) for an application that is designed to be used with remote partners. We needed our IT specialist to adjust some network settings before we could get our setup to work.
- Working around the large size of files produced by the video-capturing software we used to record participant behavior. We found we could start a new recording session at a natural break in the test session to create two smaller recordings of the test session. We also used a standard VHS video recorder on a tripod as a backup to our video-capturing software, in case our files still became too large.
- Lacking an actual system or robust prototype. We used static prototype pages and "sticky" notes for downloading, setting up, and managing the video instant-messaging service. Than we used a "look-alike" live service for actual messaging. (See Figure 3, below.)

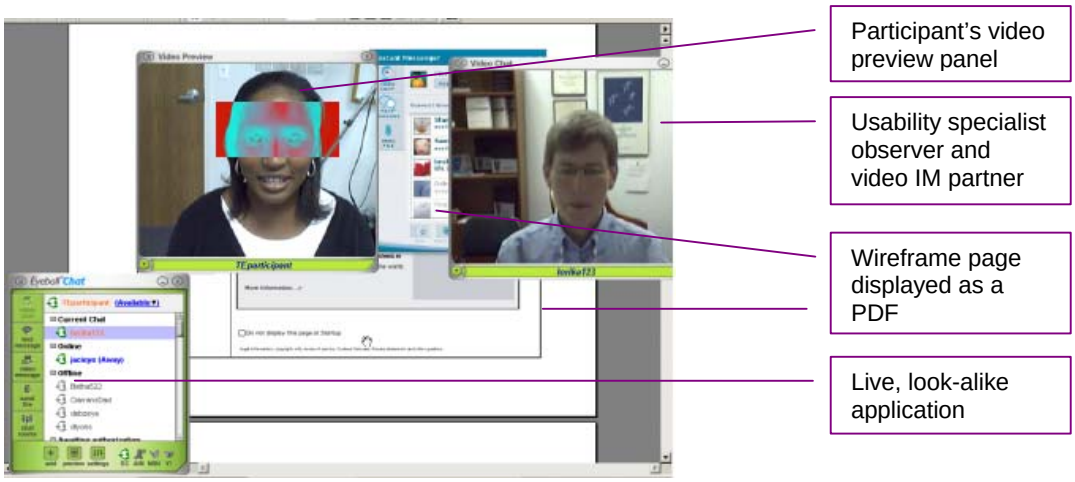
Figure 3: Prototype Wireframes, Look-Alike Service, and Sticky Note



Test Team

Because of the new technological challenges and the need for the test facilitator to manipulate the static prototype-pages to simulate the messaging software interaction, we planned from the beginning to use a two-person team. A second person served as the messaging partner, sending messages from the observation room, replying to messages from the participant, and also taking detailed notes when not acting as the messaging partner. (See Figures 4, 5, and 6, below.) The test facilitator also took detailed notes during live-messaging tasks.

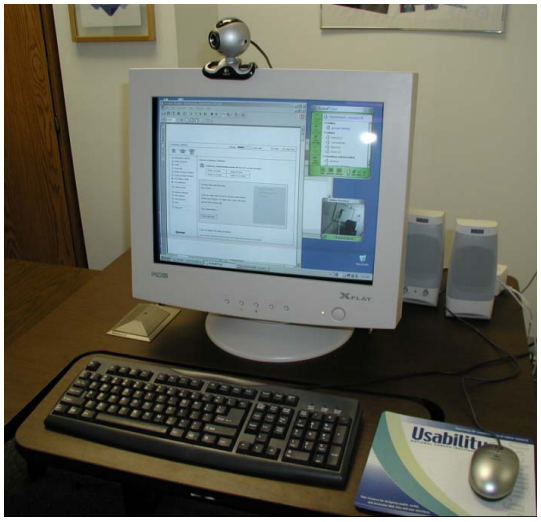
Figure 4: PC Recording of Video IM session



Technology and Equipment Setup

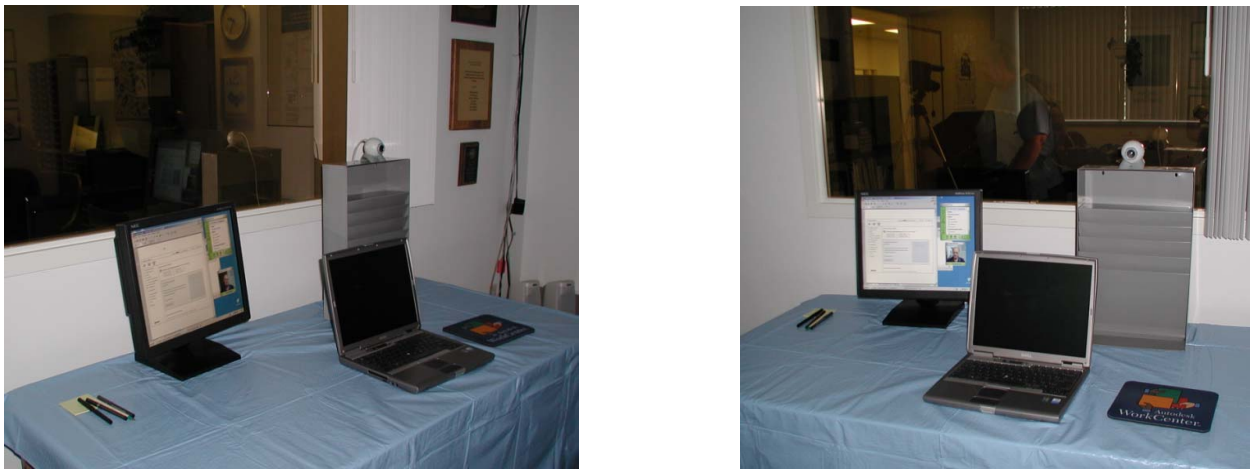
The participant sat in front of a PC connected to the Internet. External speakers, as well as a webcam with a built-in microphone, were attached to the PC, which was running a PC video-capturing program to record onscreen interaction and audio. As described earlier, for backup recording, we used a standard VHS camera mounted on a tripod aimed over the participant's shoulder at the computer screen. The first photo in Figure 5, below, is taken from the vantage point of the tripod and video camera, although the camera lens is not pointed as directly at the computer as was the video recorder.

Figure 5: Setup of Video Instant-Messaging Study — Participant and Facilitator Room



The observation room was fitted with a monitor receiving the video signal from the participant's computer and speakers receiving the audio signal from a flat, desktop microphone, plus the messaging partner's PC laptop connected to the Internet and equipped with a webcam identical to the participant's. (See Figure 6, below.)

Figure 6: Setup of Video Instant-Messaging Study — Observer/VIM Partner Room



Scenarios

Participants sent loosely scripted messages to a good friend who had "recently relocated to another city for a new job." The motivation for using video instant messaging was so that the two long-distance friends could see each other while chatting. Text messages preceded video messages to secure mutual permission and readiness for exchanging the video messages. Participants also were given small souvenirs "from a recent trip to visit mutual friends" as well as 4 X 6 photographs of the mutual friends to hold up to the webcam and show their video-messaging partner. Participants also held the camera in their hands to pan around the lab, to show their video-messaging partner their "newly reorganized computer den." The scenarios provided specific cues, spelled out in the session protocol, for the messaging partner to initiate or accept text and/or video conversations.

Conclusions

The setup that we devised worked well. We had anticipated that with this new technology, we should allow time in the schedule for testing and adjusting our proposed setup before the actual participant sessions. Therefore, we had enough time to discover and solve our internal network firewall issues and to find ways to minimize the audio echo and work around our screen/video capturing file-size limits. Our scenarios and tasks were engaging and motivated the participant to continue video chatting despite the audio challenge. The big smile on the participant's face in Figure 4, above, supports this.

Do Try This Yourself

The case studies we discussed in this paper illustrate that with careful planning and reasonably good equipment (that doesn't have to "break the bank"), IM studies are not too difficult to conduct. The authors hope that our success in overcoming challenges will encourage other usability groups to study instant messaging.

Our experience has shown that by optimally placing equipment, pre-testing setups, respecting IM etiquette and IMers' personal space, using creative, but realistic scenarios in which usability specialists can participate as well as observe, IM studies can be enjoyable and rewarding for everyone involved in the process.

REFERENCES

1. "Testing Real Work in Real Time: Evaluating Collaboration Software" (by Mary Elizabeth Raven and Kara Pernice Coyne, UPA 2000 Proceedings).

2. "Chat I: Introducing chat into business organizations: toward an instant-messaging maturity model" (by Michel Muller, Mary Elizabeth Raven, Sandra Kogan, David R. Millen, and Kenneth Carey, ACM SIGGROUP Proceedings of the Conference on Supporting Group Work).

3. Verizon Wireless, advertising campaign.

ABOUT THE AUTHORS

Deborah Hinderer Sova
Senior Usability Consultant
Tec-Ed, Inc.

Voice: 414-762-6584
Email: deobrah@teced.com

Deborah Hinderer Sova has over a dozen years of experience evaluating software and website usability for clients in a variety of industries including financial, medical, bibliographical, wireless, multimedia, enterprise software, security, and insurance. Before rejoining Tec-Ed after a three-year hiatus, she was principal consultant for Sova Consulting Group. Deborah has also worked as a user experience specialist for the Nielsen Norman Group, Optavia Corporation, American Express Financial Advisors, and Compuware Corporation. She is recognized as a participant recruiting expert, having co-authored with Jakob Nielsen an NN/g report titled "233 Tips and Tricks for Recruiting the Right Users as Participants in Usability Studies." Deborah has presented papers and participated on panels and progressions usability- and recruiting-related subjects. She has served on three UPA conference committees, most recently for the UPA 2005 Idea Market. She holds Bachelor and Master degrees from The University of Michigan.

Laurie Kantner
Senior Usability Consultant
Tec-Ed, Inc.

Voice: 734-995-1010
Fax: 734-995-1025
Email: laurie@teced.com

Laurie Kantner has defined and conducted a wide range of usability research projects as a senior staff member at Tec-Ed, Inc. Laurie has conducted numerous field usability studies using contextual inquiry and ethnographic research methods and, in the past three years, employing usability testing in the field. Studies have taken her as far as Regina, Saskatchewan, into inner-city Chicago, to the sanctum of computer company help desks and university researchers' labs, and to a rural Indiana county courthouse. Laurie also has experience designing and conducting "model projects" to transfer knowledge of usability practices and techniques to people with usability knowledge but no practical experience. Laurie is a member of the Usability Professionals' Association, the Society for Technical Communication, and ACM SIGCHI. She has a B.A. from The University of Michigan.