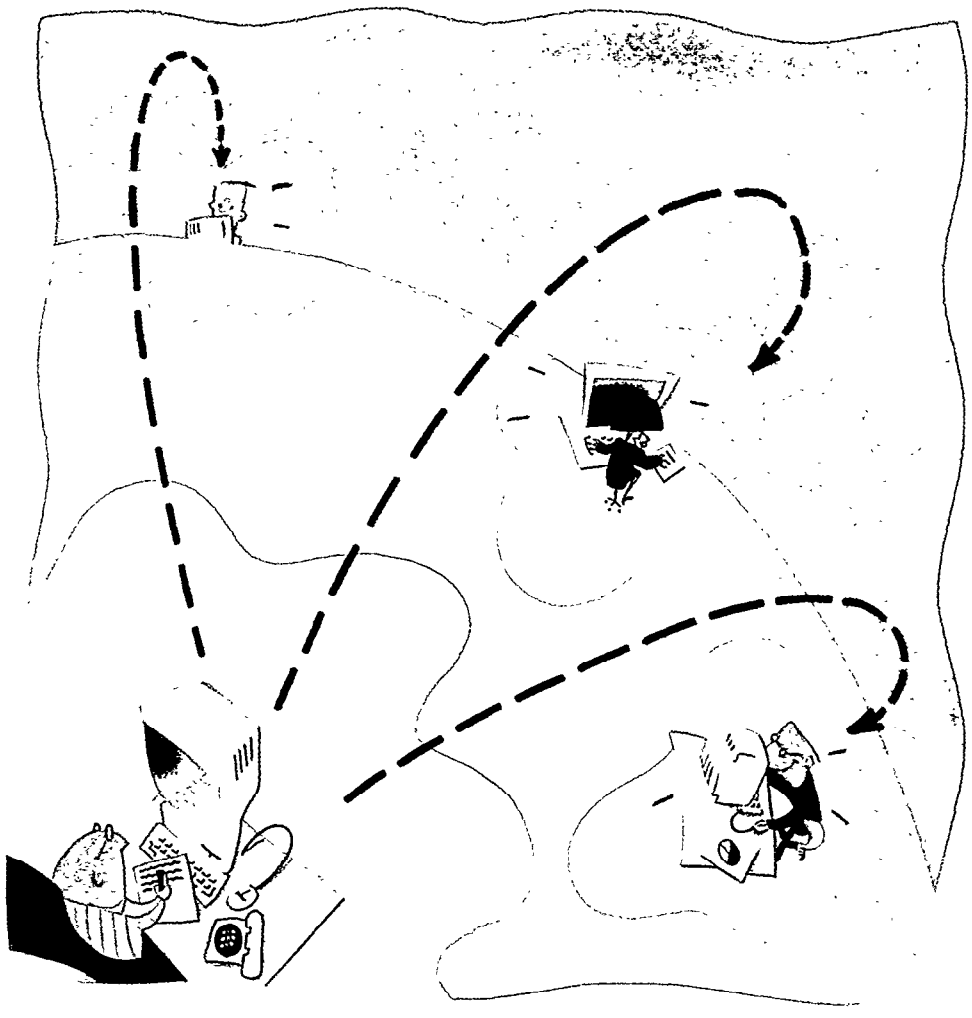
MONTY HAMMONTREE, PAUL WEILER, AND NANDINI NAYAK

Remote Usability Testing



all illustrations: Leo Comix ©

Imagine being able to collect data like static and full motion screen images, user video, and user comments, through your computer. Think about observing from your office or lab as an evaluation participant worked in their own office. What if you could sample from an intercity, interregional, or international user base without having to leave your office or lab? Thanks to recent developments in the areas of information sharing and collaborative work tools, each of these objectives is now attainable

Usability evaluators can now view computer networks and modem connections as frameworks upon which distributed usability labs can be constructed, and all network or modem accessible machines as potential windows into remote test sites. These remote test sites can range from a machine down the hall to one on the other side of the world.

The technological building blocks needed to create remote testing facilities are now available, in varying degrees, on all major computing platforms (i.e., Unix, MS Windows, Macintosh, OS/2). With computer-to-computer video conferencing tools product development teams can now both see and listen to evaluation participants sitting in front of properly equipped computers anywhere in the world. They can observe evaluation participants as they examine and annotate screen shots or documentation using shared whiteboard tools. Observers can also watch full motion images of the participant's screen from remote locations as participants use their products. By mixing and matching these technological building blocks, usability specialists can conduct a myriad of usability engineering activities remotely. The activities for which these technologies seem particularly well suited range from: participatory design exercises; to storyboard walkthroughs; to formative evaluations; to controlled experimentation; to unobtrusive field studies.

At the Sun and Hewlett Packard sites where we work, we now conduct the bulk of our testing activities remotely. Most of the remote work we have done has centered around the evaluation of interactive prototypes/applications and the evaluation of design storyboards. The following sections outline the tools, methods, and procedures, our respective groups have converged upon while working in these areas.

Interactive Prototypes/Applications:

One of the more common methods used in the testing of interactive prototypes/applications, is "thinking-aloud". During a "thinking aloud" study participants are encouraged to speak continuously about their actions and perceptions of a system as they use it to perform representative tasks. Face-to-face testing of this type conducted in a usability lab involves having users follow a set of instructions about the tasks to be performed, while usability evaluators observe. During observation, user interactions with the prototype and verbalizations are recorded, as well as non-verbal reactions (e.g., gestures, facial expressions, etc.). Usually each test session is videotaped. Frequently, an on-line logging tool is used to: facilitate note taking during the test; time specific tasks; and provide an index to the videotape for later review.

To conduct "thinking aloud" testing with users at remote test sites requires at minimum that the prototype be accessible and controllable by the remote user and that all user interaction with the prototype be visible to both the user and observers. In addition to a visual record of the user's interaction, an audio record of the user's comments must be collected.

Using a shared windowing tool (see Side-

Enabling Tools and Technologies

Usability testing can be thought of as a collaboration between users and product development teams. When seen in this light, it is a natural progression to begin looking at tools that have been developed for more general Computer-Supported Cooperative Work (CSCW) purposes as potential channels for collecting usability data. For Unix environments, Hewlett Packard's SharedX and Sun's ShowMe 2.0 are examples of these types of tools. Examples of Macintosh, MS Windows, and OS/2 products include Creative Technologies ShareView Plus for Macintosh, NCR's Telemedia Connection for Windows, and Person-to-Person/2 for OS/2. For more information about these and other related products see the September 1993 issue of Byte magazine. In general, today's CSCW tools tend to focus on one or more of the following collaborative activities.

Window/Application Sharing

With tools that support the real-time sharing of application windows between two or more workstations or terminals a user on one system can actively share an application window or a whole application with others via a network or modem link. Once sharing has occurred, both the sender and the receiver can view the shared application or window as if it were actually running on each of their respective machines.

Whiteboarding

"Shared whiteboard" applications allow multiple users to simultaneously share a common drawing/writing surface. Most of these tools allow users to pull snapshots of documents, diagrams, screens, etc... onto the shared whiteboard surface. Various participants in the sharing session can then view, point at, and annotate the shared snapshots in real-time.

Computer-based Video Conferencing

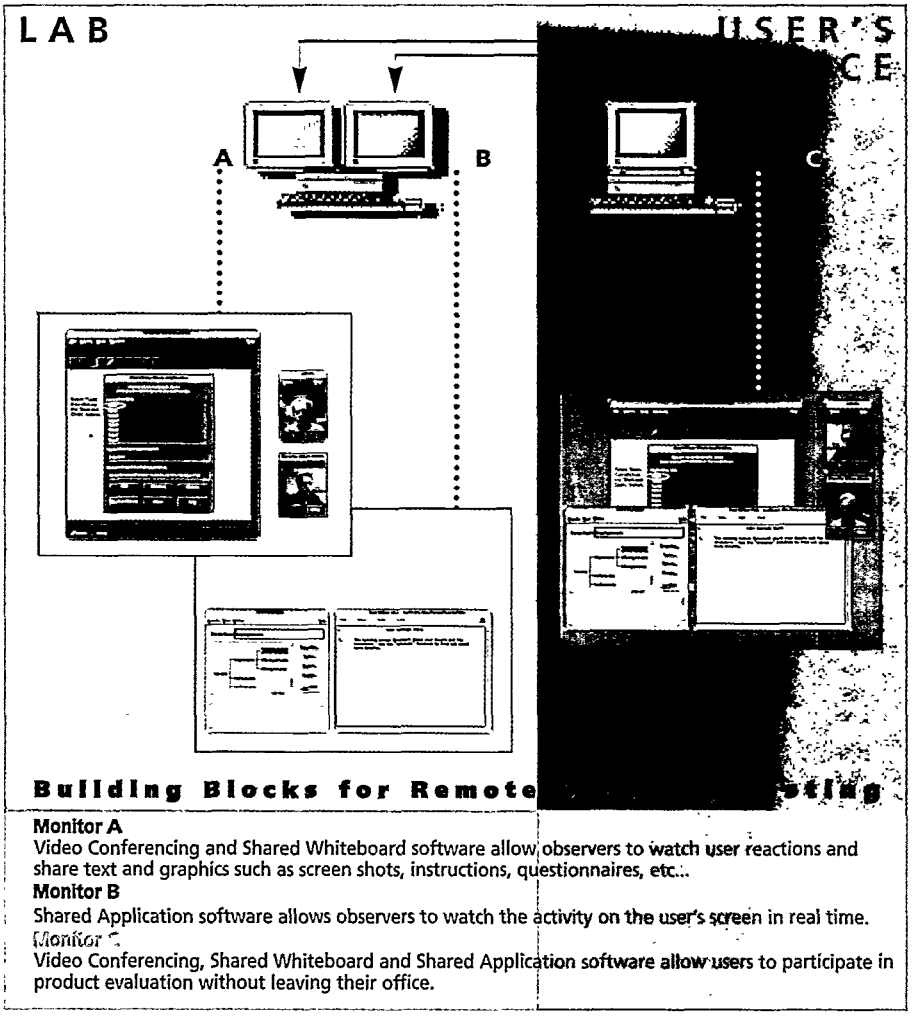
As the name indicates, computer-based video conferencing tools display live video in a window of the person or persons with whom you are conferencing. Visual cues such as gestures, facial expressions, etc... help establish rapport between conference participants and make for much richer and engaging meetings/discussions.

bar) and a telephone, it is fairly easy to put together a configuration that can support "remote thinking aloud" evaluations. Using a shared windowing tool, a prototype application can be shared from a development site to a workstation at one or more user sites. With this, as users perform assigned tasks on the shared system, all user interactions are remotely observable. A phone connection between the user and the development site makes it simple to maintain verbal communication. A videotape of the screen including the verbalizations of the user can be collected at the development site for further analysis.

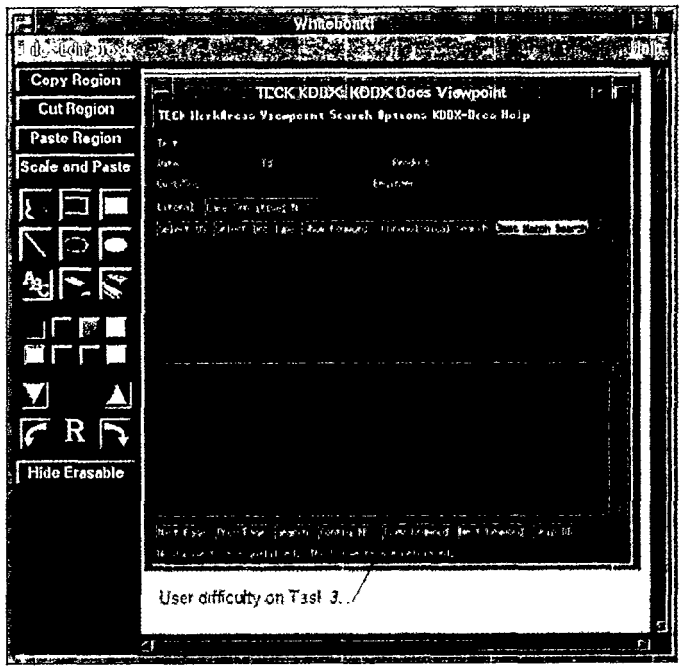
At Sun, our strategy has been to round out the afore mentioned data sources with live video of the user. Whenever possible, we augment the window sharing and phone link with a computer-to-computer video conference. This allows both participants and observers to see each other during test sessions. We feel the video link establishes a person-to-person relationship that helps build a level of rapport between the participant and the observers that is otherwise missing. We have observed that our development teams appear to be much more active and engaged in the testing process when live video of the user is provided.

Further enhancements to remote think-aloud type studies can be attained if a Shared Whiteboard is used in conjunction with the other tools mentioned. For example, if instructions and task scenarios are either faxed or emailed to the user beforehand, users may read through all the tasks before the test begins. They may in turn develop response biases that can effect their performance. To avoid such situations, a Shared Whiteboard tool can be used to present instructions for each task to users as they complete the previous task, just as users get instructions for each task on a separate page in face-to-face testing.

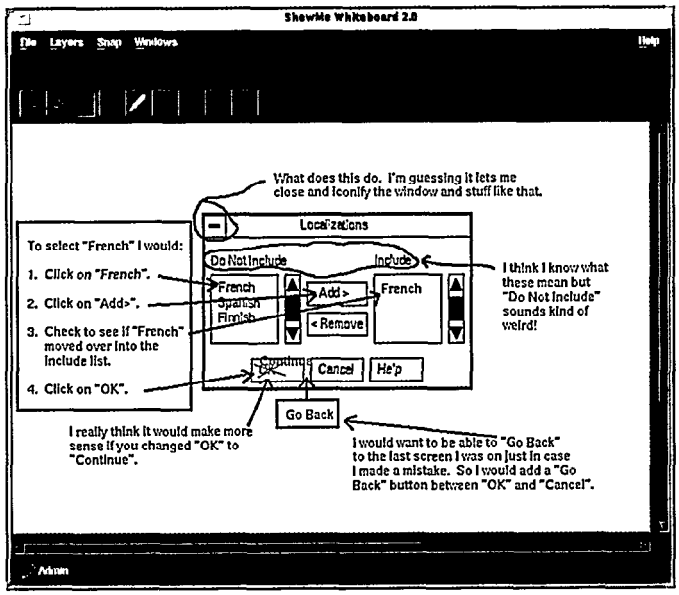
A Shared Whiteboard may also be used to capture screenshots at critical junctures during the test for later analysis. For example, if a user had problems navigating through a certain set of choices on a window, that window may be captured onto the Whiteboard. At the end of the test session, the captured screen can be shared with the user to further analyze the difficulty that was encountered and to elicit design suggestions. Both observer and user can annotate the captured screen to suggest alternative designs. These annotations can then be saved for further use. This type of event capture and replay can also be achieved through test logging tools that serve as indexes to videotapes of test sessions.



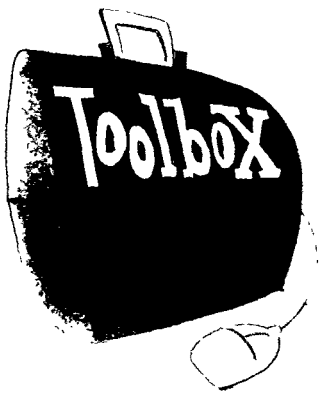
Monitor A
Video Conferencing and Shared Whiteboard software allow observers to watch user reactions and share text and graphics such as screen shots, instructions, questionnaires, etc...
Monitor B
Shared Application software allows observers to watch the activity on the user's screen in real time.
Monitor C
Video Conferencing, Shared Whiteboard and Shared Application software allow users to participate in product evaluation without leaving their office.



Captured screen being shared during debriefing.



Storyboard panel being shared during a collaborative walkthrough.



Design Storyboards

Storyboard walkthrough techniques can be used to efficiently answer many design questions For example: Are the controls on a given window, menu, or dialog box logically grouped? Are windows, dialog boxes, and controls logically sequenced? Is the sequence discernible by the user from the information on the screens alone? Are portions of the interface less intuitive than others? Are similar tasks handled consistently throughout the interface? Is the terminology used within the interface consistent with that used by end users?

The practice of methodically stepping through a design storyboard as a means of evaluating it has taken on many forms (e.g., cognitive walkthroughs, cognitive jogthroughs, collaborative walkthroughs, etc.). The primary distinguishing characteristic of Collaborative Walkthroughs is that the usability evaluator helps real users step through the interface and records their reactions along the way. This technique is most valuable early in the design process before any effort has been invested in interactive prototyping or product development. A detailed description of the Collaborative Walkthrough technique used by Randolph Bias to evaluate paper prototypes can be found in his chapter in Nielson and Mack's new Usability Inspection Methods book (John Wiley & Sons).

The first step in preparing for a Collaborative Walkthrough evaluation is to define a series of linear paths through the interface that effectively characterize the task domain that the interface should address. The second step is to develop the storyboard panels that illustrate these scenarios. Armed with on-line images of a storyboarded design, usability evaluators can easily conduct a Collaborative Walkthrough evaluation remotely using a shared whiteboard tool.

When our groups conduct Collaborative Walkthrough studies remotely, we begin by calling the test participant and setting up a speaker phone connection. We then initiate a shared whiteboard and a shared windowing conference with them. We then load the first storyboard panel onto the the shared whiteboard surface and use a shared wordprocessing window to send a description of the first task scenario. Participants are then asked to type onto the whiteboard surface the actions they would take in attempting to complete the assigned scenario. They are also asked to enter any other reactions they have toward the panel. It is important that the participants be given ample time to enter their responses before any discussion takes place and before the

next panel is loaded onto the whiteboard. After participants have entered their responses onto the shared whiteboard we ask them to discuss their responses. When participants have finished discussing their responses we tell them what the predetermined "correct" solution is. Potential usability problems are then discussed and possible solutions are entertained. After the first panel has been thoroughly discussed the next panel is presented to the participant and the process is repeated for the second panel and so forth.

Summary

Gaining access to representative test participants can be an extremely time consuming and arduous process. Often usability evaluators find themselves in situations where their target customers are highly skilled individuals whose time is in great demand. In addition, their customer bases are often widely distributed and are frequently international in scope. For example, several Hewlett Packard development teams in the United States design custom software for specific user groups located in Europe, Australia, and Asia-Pacific. In such instances no representative users are available locally and interactive user testing with representative users is only feasible if it can be performed remotely.

The recent increase in availability and data quality of tools designed to support information sharing and group collaboration has made it immediately practical to consider conducting a wide range of usability methods remotely. Present day advancements in network bandwidth, video compression, and display technologies are continuing to improve the quantity and quality of data that can be remotely transmitted. Within the near future, we believe large numbers of usability evaluators will begin to conduct the bulk of their testing activities remotely.

Monty Hammontree and Paul Weiler are human factors specialists at SunSoft in Colorado Springs, Colorado.
email: monty.hammontree@central.sun.com
paul.weiler@central.sun.com

Nandini Nayak is a human factors engineer at Hewlett Packard in Mountain View, California
email: nayakn@mayfield.hp.com