Maximizing Environmental Validity: Remote Recording of Desktop Videoconferencing

Sean Rintel

Department of Communication, University at Albany, State University of New York,
1400 Washington Ave, Albany NY 12222, USA
er8430@albany.edu

Abstract. This paper discusses the development of the technical methodology for remote recording to maximize environmental validity for a project on how novices develop familiarity with desktop videoconferencing (DVC). It is also a discussion of how the technical setup, as well as the resulting data, was useful for finding usability issues for the company that provided the DVC software.

Keywords: Desktop videoconferencing, novices, familiarity, usability, methodology, environmental validity, remote recording.

1 Introduction

Longitudinal field research on desktop videoconferencing (DVC) reveals the kinds of problems that users consider important, the way they go about solving them, and the limits of user-generated solutions. Many previous studies of DVC that stress collecting rich naturalistic data amenable to qualitative analysis have been conducted either auto-ethnographically by the researchers themselves at work [2] or conducted using novices but in fairly controlled work environments with on-site recording equipment [4, 5]. While both kinds of studies provide excellent results, they are still some distance from the high level of environmental validity that would be achieved if novice users could be recorded remotely in their own environments, on their timetable, with no extra software, and with no more effort on their part than actually videoconferencing itself. This is especially the case for home users. If such remote recording can be achieved then from both an academic and developer standpoint the resultant data should be of high quality since it minimizes the impact that the research has on the participants in the crucial moments of actually using the technology. This paper discusses the development of the technical methodology for a project exploring how novices develop familiarity with DVC. Rejected and failed solutions are discussed to provide a rationale for the chosen setup, and then the challenges and benefits of the setup are outlined for both the academic goals of the project and usability goals of the industry partner.

2 Environmental Validity

To explore how novices develop familiarity with DVC is an ethnomethodological enterprise in revealing how participants locally produce practical understandings of

J. Jacko (Ed.): Human-Computer Interaction, Part I, HCII 2007, LNCS 4550, pp. 911–920, 2007.

DVC's affordances [3] as part of accomplishing their desired social activities, and how those understandings change over time. Given that these goals are experience-based and descriptive rather than task- or technology-based and evaluative, it was decided from the outset that environmental validity was to be given primacy. In usability research, Neilson [7] defines validity as resting on whether tests measure "something of relevance to usability of real products in real use outside the laboratory." Ethnomethodological usability research does not depend on measurement, but its concept of validity could be said to share Neilson's emphasis on relevance in situated reality [6, 8].

For this project, I use the term "environmental validity" to mean what most sociological research refers to as "ecological validity", that is, as a concern that a controlled research situation should approximate the real-life situation that is under investigation. I prefer to use the term environmental validity both to avoid conflict with the Brunswickian [1] definition of ecological validity and as a way of stressing that the physical environment in which DVC takes place is likely to matter a great deal to how people use it. As such, a primary concern of this project was to approximate the real–life experience of novices trying out DVC for maintaining long-distance personal relationships and to capture that experience as richly as possible. 12 pairs were to try DVC for two to three months. A combination of observational and interview data was decided upon, with the observational data having primacy. This meant enabling DVC in the participants' own homes and recording what occurred for later analysis. The observational data collection objective was an unbroken longitudinal record of every DVC interaction engaged in by a pair.

3 Environmental Validity on a Shoestring Budget

3.1 Rejected Local Recording Solutions

Studies on DVC at work frequently use recordings of both on-screen action and action in the physical environment [2, 4, 5]. For this project on home DVC, however, the latter had to be ruled out. Not only were resources too limited for physical environment recording, but the physical and social intrusion that would have resulted from placing suitable recording equipment in homes would have undermined the naturalness of the trials, not to mention causing extreme ethical concerns. On-screen DVC action, then, was to be the crux of the observational data.

To conduct moment-by-moment microanalysis of interactions by pairs via DVC it is critical to record the synchronized video and audio of each pair member as they experienced it, and that those recordings are synchronized with one another such that the what the researcher experiences after the fact closely approximates the experience of the participants during the interaction. There were, then, two linked issues. First, how the DVC video and audio for each participant could be accessed and recorded, which impacted directly on environmental validity. Second, how the recordings of separate participants could be synchronized, which was affected somewhat by the requirements of environmental validity. The intrusion and resource factors that led to ruling out physical environment recording also impacted upon the methods of accessing and recording the on-screen DVC action.

Methods for recording DVC with apparatus in participants' homes were rejected for several reasons. The simplest rejected method was a VCR connected to each participant's computer that the participants would start and stop for every DVC event. This was rejected primarily due to intrusion problems. Requiring participants to record themselves is highly unnatural, emphasizing their awareness of the unreality of the situation and the knowledge that there is a persistent record of what would ordinarily be ephemeral. The VCR would have to be fitted into the participants' environments in such a way as to be unobtrusive but still easily accessible, as would an unknown quantity of blank video-cassettes, and recorded video-cassettes would have to be changed, labeled and stored. Participants would also have had to be trained to use and troubleshoot the VCR (e.g. how to reset the clock if the power went off, since a non-functioning clock often prevents any action until reset). Further complicating such a method was the possibility that participants might be using laptop computers, adding the need to frequently un-plug and re-plug the VCR into the computer. As it turned out, over 80% of participants in the study used laptops, so this could have been a very real problem. Even this 'simple' solution would allow for both deliberate and accidental lapses in recording, problematizing the longitudinal record.

More complex, automated recording systems, using off-the shelf or purpose built digital video recorders, were also rejected for similar reasons. While the physical impact of equipment access and replacing and storing tapes would have been eliminated, the laptop issue would still have been a factor. And, of course, designing and supplying a bespoke easily un-and-re-pluggable automated recording system was well beyond the resources of the project. The possibility of using automatic screen-capture software loaded directly on participants' computer (e.g. Camtasia) was also tested, but two more problems surfaced. First, the size of captured video files was overwhelming. Participant computers would have had to have vastly upgraded hard drives just to fit the files. Further, in tests on a laptop and desktop of average power, the rendering of the temporary video file for even a twenty minute conversation either crashed the computer or took an inordinately long time. Clearly these issues would have directly decreased the realism of the situation for the participants. Local screen captures were also rejected by the researcher because tests found that after an indeterminate period of time above around five minutes, the video and audio of the recordings would desynchronize, rendering the recordings unusable.

Even had they been feasible, any form of local recording was also determined to add a layer of complexity to working up the data as analyzable recordings because of the need for synchronizing individual records. While both VCR and digital video recordings can be synchronized using various forms of time-codes, doing so was not worth the effort given both the environmental validity and practical problems that these systems presented. Having rejected local recording systems, remote recording was the only remaining solution. But coming up with an affordable, unobtrusive, and naturalistic system was also a challenge.

3.2 Rejected Remote Recording Solutions

Remote usability testing solutions have flourished in the last few years (REF). Most purpose-built remote usability software (e.g. Morae) and remote desktop sharing solutions (e.g. RealVNC, Microsoft Remote Desktop, Citrix, WebEx) work on the same principle: users load and run the remote sharing software locally, and then all on-screen activity is uploaded in real time to the remote research location. This provides the richest possible picture of the use of the application in question, as well as how users multi-task. The obvious benefit of such a solution is that environmental validity of the project is increased because recording becomes essentially transparent to the participants: There is no physical impact on the participants' environments and, since these systems can be run automatically and silently on computer startup, also greatly reduced impact on the actions that participants need to take per DVC event. WebEx and Morae were beyond the resources of this project, but both RealVNC and Remote Desktop (built into Windows XP Pro) were affordable enough to be possible data delivery systems. In theory these remote sharing solutions seemed ideal for this project, but practical obstacles led to their rejection.

It was immediately found that RealVNC, in common with most frame-buffered desktop sharing applications, does not support transmission of the remote computer's audio, making it unsuitable for this project. Desktop sharing applications based on the Remote Desktop Protocol (RDP), however, such as the Remote Desktop built into Windows XP Pro, do support audio transmission and improved video transmission, and thus tests were conducted using Remote Desktop in Windows XP Pro.

The biggest barrier to using remote sharing was participant upload bandwidth limitations. Tests showed that the average capped upload limit of around 300kbps of home broadband connections was too limited to support both DVC and a real-time remote desktop connection with full fidelity video and audio. Those DVC applications that did not freeze had heavily degraded video and audio streams. Indeed, during the project it was found that this 300kbps average upload rate was on occasion hard-pressed to transmit the DVC upload stream itself, given that many users were (a) sharing their network connections and (b) despite advice to the contrary, running other applications which used upload bandwidth. While home users continue to have severely limited upload bandwidths the potential for using RDP-based testing of bandwidth-heavy home applications seems to be limited. There is, however, hope for improvement. The remote sharing solution also had the same synchronization problem as local recording solutions: Individual pair members recordings would have had to be combined and synchronized. However, unlike the local recording solution, had the remote sharing solutions worked to adequately deliver the DVC streams, this would certainly have been worth the effort.

The next remote recording solution envisaged was directly tapping into the streams of DVC systems' servers at their source. This would have split the participant DVC streams as they occur and direct the recorded versions into recording apparatus at the server location. No software other than the DVC application itself would be in operation, leaving participants' upload bandwidth streams untouched. Clearly this solution would have improved upon the remote sharing solution because it would be completely transparent to users, providing the highest level of environmental validity possible. The old story of resource limitation reared its head in two ways to prevent this solution. Not only did the project lack the purely financial resources to run a dedicated DVC server but also the resources to develop software to enable splitting and copying of the DVC streams. Had these resources been available, this would have been an excellent solution.

3.3 Remote Recording of Participants Using a Multi-party Bridge

Having eliminated all local and most remote recording methods due to resource scarcity and validity problems, the only solution remaining was to have participants interact in a small multi-party situation. A pair of interactants and a third party who does not interact are just a form of small group, and thus so long as a group DVC room/channel could be limited to just the desired interactants, and participants consented to being recorded, this solution would provide the desired access. Indeed, not only would access be possible, the access would be largely the same for researcher as the participants, and it would have the virtue of allowing one recording to capture all the interaction, removing the problem of synchronizing separate recordings. Although one extra step would be required above regular point-to-point DVC—pairs would have to log into a multi-party bridge before seeing their partner and then talking—it would have a very limited impact upon environmental validity, certainly much less than other methods.

However, there was an important environmental validity problem to be solved for remote recording using a multi-party bridge. Most multi-party DVC video window interfaces are designed to indicate all participants in a room/channel to enhance the sense of group cohesion. Apple iChatAV and SightSpeed are just two of many DVC services which display all participants within a single video window. Non-participating members are marked by very obvious place-holders if no camera is attached (mono-colored areas or iconic representations) or a view of an empty physical area if a camera is attached (See Figure 1).

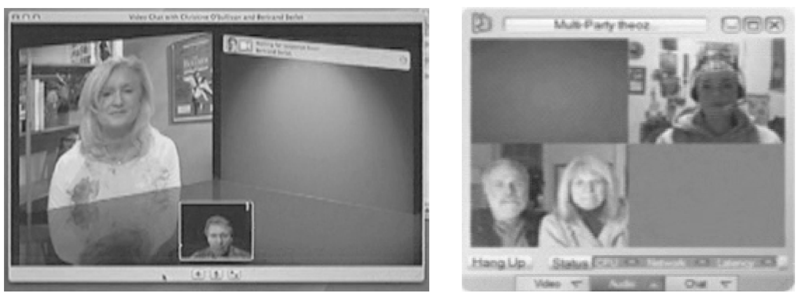

Fig. 1. Displays from Apple iChatAV and SightSpeed in which non-participating members are marked by overly obvious place holders

While this is useful when the group context needs to be foregrounded, it was a drawback for this approach to remote recording since a placeholder would immediately and constantly alert participants to the fact of their being recorded, significantly reducing sense that they were having a private dyadic interaction. Pilot testing using SightSpeed indicated that participants would discuss the placeholder, thus it was critical to find a DVC application that minimized visibility of the researcher.

Extensive trials led to the discovery that two DVC services took an alternate approach to multi-party participant video display. Regardless of the number of participants, both Wave Three Inc.'s Session Communication software and iVisit displayed all participants' video in individual floating windows that could be resized and positioned anywhere on the desktop. This allowed a group of three to appear as a group of two, since the participants could choose to start only the video windows they wanted and position them anywhere they wanted. Session was chosen for its superior quality video and audio in tests. Figure 2 shows images of Session windows open on participant desktop and the server desktop.

Fig. 2. Session video windows as they appeared on a participant's desktop and on the server desktop (participant desktop image is a mockup based on Session's display capabilities and participant descriptions of video window placement)

Session also had several features which minimized visibility of the researcher. First, its primary call window was small and easily hidden, giving primacy to the participant video windows. Once a conversation was underway, participants were not constantly confronted with evidence of the recording beyond a single name in their contact list, which most participants reported placing away from the video windows or minimizing so as to not be visible at all. Second, when participants logged into their multi-party bridge, although they would see the researcher's computer on their contact list along with the name of their conversational partner, no video would start automatically. Since starting video required participants to click on a play button next to a contact's name, participants were told to just start their partner's video and ignore the researcher's contact entirely. This quickly led to participants simply logging on to the bridge, starting their partner's video, and talking. While participants did report awareness of being recorded, none reported feelings of intrusion during interactions.

Having solved the access issue and associated environmental validity problem, all that remained was a suitable recording system. Four computers were permanently connected to four separate bridges 24/7 as non-participating contacts. Each recording computer output its video and audio to a video cassette recorder (VCR) with a timer set to record for 9 hours. Figure 3 shows a diagram of the final multi-party bridge setup. With daily tape changes this remote recording system ran seven days a week for nine months.

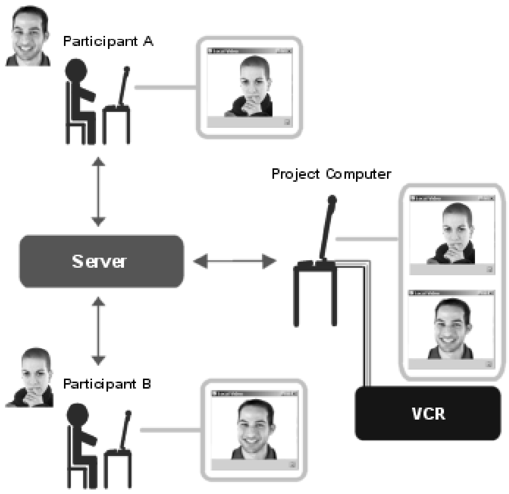

Fig. 3. Multi-party bridge remote recording setup

Had more resources been available, an A/V distribution box and three VCRs per computer could have been used for full 24/7 recording. As it was, an automation script (to be discussed shortly) logged all Session activity—logon, logoff, participant connection, errors etc.—which not only helped pinpointing the times when participants actually began videoconferences, but also occasions when participants logged in outside of the recorded times and also with the diagnosis of technical problems. However, unplanned contingencies lead to a number of challenges, which along with being useful for refining the setup, also provided usability data in and of themselves.

4 Challenges

Automation of the recording servers was necessary to allow participants to interact on their own schedules. This meant running automation scripts to control Session. The two major areas of automation required were video display and bridge connection. Automating video display was crucial. As was discussed above, one of the features of Session was that in multi-party bridge situations, no participant video was started automatically. Wave Three's rationale for manual video startup in bridges is that it prevents participants' bandwidth being swamped, which is reasonable and participants reported having no problems starting video manually but most would have preferred set automatic startup. Nevertheless, the need for manual video startup contributed to the environmental validity for this method, since the non-participating researcher's computer video was not started for participants. Unfortunately, this also meant that, if left unattended, the recording computers would record only participant audio. A script was needed to automatically detect when users had connected to the bridge and then simulate the clicking of the play button next to their contact names. Automating bridge connection was also crucial to ensuring that participant interactions could be recorded at any time. A script was needed to check that the bridge connection was constantly up, and to reconnect if it went down for any reason.

Auto It v3 was used to create scripts for the recording computers running Microsoft Windows. Like most scripting applications, Auto It scripts generally rely on being able to identify and interact with application controls based on unique textual identifiers from the Windows API. However, since Session is a Java application wrapped in the standard Microsoft Windows skin, Auto It could not access the application controls directly. The solution was to use a combination of indirect methods: watching for the titles of windows and triangulated sets of color changes on the Session interface and desktop to determine what actions to take at any given time. While this inelegant kludge solution turned out to be effective, in practice it took a little time to smooth out. This is because for the first few months of the project unexpected color display events prevented script triggers, occasionally preventing video display or bridge connection.

Some unexpected color display events were purely technical and occurred only at the recording computer. These were primarily unexpectedly opening system or applications windows—especially automated application update windows—covering parts of the Session window or desktop that the script was watching. The quantity and variety of unexpectedly opening windows was surprising, and it took some time to ensure that as many unexpected window openings could be suppressed as possible. Other unexpected color events occurred because of participant actions. An early version of the script expected participants to use assigned user names and searched for color on parts of those predefined names. This was optimistic. When one member of an early participant pair did not use an assigned username (which was intended to prevent identification as well as enabling automation), the script was not triggered and no video was recorded for that participant.

Automated bridge connection and reconnection were initially included in the scripts as a time-saving tool rather than a necessity, as no bridge disconnection ever occurred during testing. However, as the first round of data collection got underway it became apparent that the Session servers were not designed to support users staying logged in to bridges 24/7 as the recording computers were. The first three month data collection round experienced bridges downtime on 24 occasions: usually for a day and a night, occasionally over a weekend, and once for 4 nights/3 days. After discussions about fixes led to daily reboots of the Session services, the second data collection period of four months experienced only 7 downtimes: usually just overnight, but once for 3 nights/2 days. The bridges could only be restarted by Wave Three support personnel during regular US Pacific Time working hours. While the researcher was checking the bridges several times a day and could notify Wave Three when problems were noticed in these timeframes, some bridge downtime occasions prevented participants from being able to videoconference on demand. Staying permanently logged in to a bridge is admittedly unusual—bridges are designed more for temporary use—but at the time even Wave Three Inc. were not aware that this kind of login would stress the otherwise fairly solid server system. Wave Three is currently diagnosing the cause of these bridge downtimes, an issue which might not have been discovered in ordinary usage. There is, of course, some irony that to provide environmental validity an unusual use of the servers had to be instituted which then caused some problems.

5 Results

Given the fairly long road to the technical methodology outlined above, and the challenges faced in implementing it, was the resulting data worth the effort? I believe that it was. Although a combination of technical and scheduling issues with users during setup had to be ironed out to get them talking frequently enough to provide adequate data (which may be the subject of another paper), once most pairs moved into actually using DVC, it quickly became apparent that the comfort level of using DVC in their own environments was quite significant. Indeed, the first pair to complete the study was so comfortable using DVC late at night in their bedrooms that half of their hour-plus conversations would end with one or both falling asleep! Falling asleep 'with each other' after a long late-night conversation was a pattern that this pair carried over from their pre-trial behaviors during mobile telephone conversations.

Although that particular manifestation of comfort was unique, common to all pairs was the spontaneous display and employment of physical objects from the local physical environment. This, of course, would be unlikely in a laboratory environment, as would the personal nature of the items have been in a controlled work environment. Perhaps more interestingly, the real clue to high environmental validity is that displays of objects or behavior by one pair member could be reciprocated by the other. For example, at one point in a conversation one participant decided that he would put on sunglasses. "We're wearing sunglasses now?" asked his pair member, rummaging around her room to find her own pair, which she then put on (Figure 4).

Fig. 4. Display and reciprocated display of personal items from the local environment

Also common (and reciprocated) were numerous incidents of talking while folding laundry (Figure 5), eating, pulling faces (Figure 5), taboo gestures (e.g. 'the finger') and discussion of bodily functions.

Fig. 5. Participants show comfort with the recording situation: Doing laundry and pulling faces

What is important about all of the behavior above is not, of course, that it involves particularly significant incidents. Rather it demonstrates that despite most participants reporting being generally aware of being recorded, the environmentally validity of the project situation was high enough that the participants felt comfortable enough to do frequently do ‘as they would’ instead of only ‘as they ought’.

6 Conclusions

This project on how novices develop familiarity with DVC was premised on the high environmental validity of the trial experience for the novices. Using the multi-party bridges of the Session software allowed the researcher to have virtually the same access to the interactions as that experienced by the participants while not burdening them with equipment in their local environment or extra software on their computers to use resources or bandwidth. Further, the fact that Session displayed user video in separate windows and did not automatically start any contacts’ video automatically meant that the recording was almost transparent to the participants. For the industry partner, an interesting side effect of the project connecting to the bridges 24/7 was to turn up a problem with the Session bridge server system that had not been previously found and could prevent on-demand DVC: a critical usability issue. As far as the academic project itself is concerned, the comfort level displayed by the participants bears out the contention that the technical methodology approximated real-world usage enough to make valid claims. Thus while this solution came about as a response to a shoestring budget and technical limitations, and only after rejecting many other solutions, it turned out the resulting technical methodology provided real benefits for environmental validity.

References

1. Brunswik, E.: Perception and the Representative Design of Psychological Experiments, 2nd edn. University of California Press, Berkeley CA (1956)
2. Dourish, P., Adler, A., Bellotti, V., Henderson, A.: Your Place or Mine? Learning From Long-Term Use of Audio-Video Communication. Computer-Supported Cooperative Work. 5(1), 33–62 (1996)
3. Gibson, J.J.: The Ecological Approach to Visual Perception. Lawrence Erlbaum, Hillsdale NJ (1979)
4. Heath, C., Luff, P.: Technology in Action. Cambridge University Press, Cambridge UK (2001)
5. Heath, C., Luff, P.: Disembodied Conduct: Communication Through Video in a Multi-Media Office Environment. In: Proc. CHI 1991 pp. 99–103 (1991)
6. Garfinkel, H.: Studies in Ethnomethodology. Prentice-Hall, Englewood Cliffs NJ (1967)
7. Nielson, J.: Usability Engineering. Academic Press, San Diego CA (1993)
8. Suchman, L.A.: Plans and Situated Actions: The Problem of Human-Machine Communications. Cambridge University Press, Cambridge UK (1987)